DOCUMENT RESUME

ED 209 677 CS 206 645

AUTHOR	Witte, Stephen P.; Faigley, Lester
TITLE	A Comparison of Analytic and Synthetic Approaches to the Teaching of College Writing.
INSTITUTION	Texas Univ., Austin. Dept. of English.
REPORT NO	TWRG-RR-1
PUB DATE	Aug 81
NOTE	472p.
EDRS PRICE	MF01/PC19 Plus Postage.
DESCRIPTORS	College English; College Freshmen; Program Effectiveness; *Writing Evaluation; *Writing Instruction; *Writing Research
IDENTIFIERS	*Freshman Composition; *University of Texas Austin

ABSTRACT

This report provides information drawn from a study that compared two components of the freshman English program at the University of Texas at Austin. The first chapter of the report contains background material and describes the two course components--the analytic option, which emphasizes writing skills in the context of complete pieces of writing, and the synthetic option, which works from individual discourse parts to paragraphs to complete essays. The second chapter discusses the research design and procedures, while the third chapter analyzes some features of high and low quality student essays. Chapters four and five offer comparisons of essays written by students in the analytic option with those written by students in the synthetic option. Chapter six discusses the influence of writing apprehension on student performance in freshman composition, and chapter seven outlines the predictors of writing quality and course grades. The final chapter places the study in the larger context of writing research conducted at other universities and presents a summary of the study's findings, its limitations, and its implications for teaching and research. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209677

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TEXAS
WRITING
RESEARCH
GROUP

A COMPARISON OF ANALYTIC AND SYNTHETIC APPROACHES TO THE TEACHING OF COLLEGE WRITING

TWRG RESEARCH REPORT #1

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Stephen P. Witte
Lester Faigley
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Authors: Stephen P. Witte
Lester Faigley

August 1981
The University of Texas at Austin
English Department—PAR 110
Austin, TX 78712

CS206645

A COMPARISON OF ANALYTIC AND SYNTHETIC APPROACHES TO THE TEACHING OF COLLEGE WRITING

by

Stephen P. Witte

and

Lester Faigley

The University of Texas at Austin

August 1981

copyright © 1981 by Stephen P. Witte and Lester Faigley.

Preface

This tome presents in some detail a comparative evaluation of two components of the Freshman English Program at the University of Texas at Austin. The evaluation focuses on the Freshman English Program as it existed in the Fall of 1978. The evaluation was begun in August of 1978 by Stephen Witte with the assistance of James L. Kinneavy, Director of Freshman English, and Thomas Cameron, an Assistant Instructor in the Department of English. The analyses reported here were completed by Witte and Lester Faigley, who joined the University of Texas faculty in September of 1979, during the 1979-1980 academic year.

The authors wish to thank a number of persons who contributed directly to the study in various ways. For Jim Kinneavy's advice on the design of the study and for his making available to us some of the resources of the Freshman English Office, we are most grateful. For Tom Cameron's careful attention to the many details associated with collecting the large amounts of data required for the study, we are especially thankful. Without Kinneavy's and Cameron's contributions this study could not have been completed.

One of the major difficulties encountered in completing a study of the magnitude of this one is organizing the various pieces of data which were collected from each student. Merry Eginiore, Paula Marks, Paul Meyer, and Tom Cameron provided valuable assistance in that area.

Many of the analyses reported in the following pages depend on the holistic ratings given to the four sets of student essays collected. A number of persons served as raters, giving of their time and expertise as teachers of college writing. We are grateful to them for their expert help, patience, and good humor. They are: Bruce Edwards, Camy Cunningham, Paula Marks, Paul Meyer, David McMurrey, Richard Hart, Michael Morgan, Alice Korach, Mark Akers, Kathryn Staples, Miriam Hamilton, Rebecca Beal, Susan Blalock, Christine Cowden, Sue Chisholm, Carol Edkins, and Tom Cameron.

Gay Marie Logsdon, Richard Hart, Thomas Cameron, Wayne Bowen, and Marianna DiPaolo assisted us most capably with the syntactic analyses of the student essays after the essays had been rated holistically. We are grateful for their having performed most capably a very tedious task.

Our typists--Rebecca King, Karen Lunos, Wayne Bowen, and Chris Benton--deserve more than this short note of thanks for having rendered our imperfect handwriting into legible typed pages.

A number of fellow professionals offered sound advice at various points during the study. In addition to Jim Kinneavy, Ellen Nold (Stanford), Bob Kline (Curriculum and Instruction, University of Texas), and Bill Davis (late of the University of Wisconsin-Madison) made helpful suggestions about the design of the study. Paul Kelly, Susan Hereford, Wayne Fisher, and Earl Jennings (all of the University of Texas Measurement and Evaluation Center) offered good advice on some of the methods used during the analysis.

We are grateful to Bob Kline for having read parts of the

report which follows. His comments were helpful at various stages during the writing process. We are equally grateful to Maxine Hairston for having read the typescript for the Texas Writing Research Group and for approving it for publication by that body and to Cindy Selfe for seeing that the manuscript was printed.

Finally and most importantly, we would like to thank the students and the teachers of the classes we evaluated. Without their cooperation we could not have done anything. We cannot, of course, thank the students individually, although their cooperation and good faith were essential. We do, however, extend our thanks to the teachers who endured through the ordeal of collecting data and gave us their appraisals of their courses and the evaluation project. These teachers were Noreen Kornman, Bruce Gentry, Mary Jean Gross, John Orth, Martha King, Jim McNutt, Debbie Hendrixson, Steve Harris, and Margaret Mahoney.

Stephen P. Witte
July, 1981

TABLE OF CONTENTS

PREFACE iii

CHAPTER 1: GENERAL INTRODUCTION 1

1.1 Background of the Present Study 1
1.2 The Two Major E 306 Options 2
    The Analytic Option 2
    The Synthetic Option 4
1.3 Use of Data Collected During the Evaluation 6

CHAPTER 2: RESEARCH DESIGN AND PROCEDURES 7

2.1 The Designs of Ten Comparisons 7
    The Five Principal Comparisons 8
    Sample Population and Control of Teacher Variable in the Five Principal Comparisons 9
    The Five Derivative Comparisons 10
2.2 Pretest and Posttest Data 11
    Measures 12
    The Writing Topics 12
    Administration of the Test Instruments 16
    Sample Selection 18
2.3 Holistic Evaluation Procedures 19
    Preparing the Essays for Holistic Evaluation . . 19
    Single-Impression Holistic Evaluation 20
    The Distribution of Holistic Scores by Scoring Category 24
2.4 Syntactic Analyses of the E 306 Essays 27

CHAPTER 3: SOME FEATURES OF "HIGH" AND "LOW" QUALITY E 306 ESSAYS 29

3.1 Error Analysis of E 306 Essays 32
    Punctuation 33
    Grammatical Errors 35
    Spelling Errors 36
    Comparison of "High" and "Low" Quality Essays . . 36

3.2 Syntactic Analysis of E 306 Essays 44
Mean T-Unit Length 47
Mean Clause Length 51
Subordination Index 52
Coordination-Within-Sentence Index 53
Mean Sentence Length 53
Free Modifiers 54
Comparison of "High" and "Low" Quality Essays 58

3.3 The Occurrence of First-Person and Second-Person Pronouns in "High" and "Low" Quality Essays . . . 69

3.4 Text Cohesion 73
Cohesion Through Reference 80
Cohesion Through Substitution 85
Cohesion Through Ellipsis 88
Cohesion Through Conjunction 91
Lexical Cohesion 95
Comparison of "High" and "Low" Quality Essays 104

3.5 Errors, Syntax, Personal Pronouns, and Cohesion in E 306 Essays: A Conclusion 124

CHAPTER 4: E 306 CHANGE SCORES AND FIVE DERIVATIVE COMPARISONS 127

4.1 Changes in Writing and Writing-Related Behaviors . 128
Changes Identified by the Pre- and Posttest Essays 129
Changes Identified by Scores on Objective Tests 137

4.2 The Five Derivative Comparisons 145
Derivative Comparison 1: All "Analytic-Conventional" Classes Compared with All "Synthetic-Conventional" Classes 147
Derivative Comparison 2: All "Analytic-Conventional" Classes Compared with All "Synthetic-Tutorial" Classes 154
Derivative Comparison 3: All "Conventional" Classes Compared with All "Tutorial" Classes 160
Derivative Comparison 4: All "Analytic" Classes Compared with All "Synthetic" Classes 165

Derivative Comparison 5: All "Analytic-Conventional" Classes Compared with All "Analytic-Tutorial" Classes 173

4.3 Chapter 4: A Summary 178

CHAPTER 5: THE FIVE PRINCIPAL COMPARISONS 183

5.1 Review of the Five Principal Comparisons 183

5.2 The Result of the Five Principal Comparisons 185

Principal Comparison 1: "Analytic-Conventional" Students (N=18) Compared with "Synthetic-Laboratory/Tutorial" Students (N=18) 185

Principal Comparison 2: "Analytic-Conventional" Students (N-18) Compared with "Synthetic-Conventional" Students (N=18). 195

Principal Comparison 3: "Analytic-Conventional" Students (N=18) Compared with "Analytic-Syllabus-Conventional" Students (N-18) . . . 205

Principal Comparison 4: "Analytic-Conventional" Students (N=18) Compared with "Analytic-Tutorial" Students (N=18) 217

Principal Comparison 5: "Analytic-Tutorial" Students (N=18) Compared with "Synthetic-Tutorial" Students (N=18) 228

5.3 The Five Derivative Comparisons: A Summary 237

CHAPTER 6: THE INFLUENCE OF WRITING APPREHENSION ON STUDENT PERFORMANCE IN E 306 243

6.1 Summary of Previous Research 243

6.2 E 306 Writers and Writing Apprehension 244

Method 245

Results 246

Discussion 247

6.3 Conclusion 249

CHAPTER 7: PREDICTORS OF WRITING QUALITY AND COURSE GRADE IN E 305 251

7.1 Summary of Previous Research 251

7.2 Predictors of Writing Quality in E 306 253

7.3 Predictors of Course Grades in E 306 256

7.4 Discussion of the Results 258

CHAPTER 8: THE PRESENT STUDY IN THE LARGER CONTEXT 265

8.1 The Larger Context of the Present Study 265

The University of Northern Iowa Study 266

The University of California at San Diego Study 268

The Miami University Study 269

The University of North Dakota Study 272

8.2 Summary of the Major Findings of the Present Study 278

8.3 The Limitations of the Present Study 284

8.4 Implications for Teaching and Research 288

FIGURES AND TABLES 293

BIBLIOGRAPHY OF WORKS CITED 403

CHAPTER 1

GENERAL INTRODUCTION

1.1 Background of the Present Study

The present study was undertaken in the fall of 1978 for the specific purpose of comparing the relative effectiveness of two curricular and instructional options for acquiring credit in English 306, the introductory course in rhetoric and composition at the University of Texas at Austin. We call these two major options the analytic option and the synthetic option. They will be described in detail below.

During the Fall of 1978, the analytic option served as the introductory composition course for nearly 3,300 UT freshmen and the synthetic option served 130 freshmen. These students had not received exemption credit by examination for E 306. The unequal distribution of students within the two options had continued for approximately four years prior to the study. Students not only avoided enrolling in the synthetic courses, but thirty percent of the students who did enroll in the synthetic tutorial courses withdrew from them, an attrition rate twice as high as for the analytic courses.

Few data existed indicating whether the option attracting the fewer students was equal to, worse, or better than the analytic option. Faced each year with the higher attrition rate and the higher cost per student credit hour of the synthetic option, Professor James Kinneavy, the Director of Freshman English, suggested a systematic comparison of the two options. Thus in August of 1978, Prof. Kinneavy, Prof. Stephen Witte, and Thomas Cameron, a doctoral student in English Education, wrote a research proposal calling for the comparative evaluation which is the subject of this report. During the analysis stage of the study, Prof. Lester Faigley, who joined the UT faculty in fall of 1979, worked with Prof. Witte on the project.

1.2 The Two Major E 306 Options

The Analytic Option. Several fundamental differences obtained between the two major E 306 options during the Fall Semester of 1978. We call the first option the analytic option because of its emphasis upon teaching writing skills in the context of complete pieces of writing rather than in isolation. The option served the larger number of UT freshmen and was taught in a conventional classroom setting with lecture-discussion as the principal instructional medium, although, of course, student conferences were a regular part of the instructional method. The rhetorical framework for the analytic option was provided in part by James Kinneavy's A Theory of Discourse (1971) and presented through a syllabus written primarily by Sue Rodi (1978) of the UT English Department. The underlying theory, with its emphasis on whole discourses, provided the basis for a holistic approach to the teaching of writing, and thus attended to the development of writing skills in the context of three purposes of written discourse--the expressive purpose, the persuasive purpose and referential purpose--as realized through four discourse modes--description, classification, narration, and evaluation. Those interested in the derivation and application of several principles underlying the analytic curriculum should consult Kinneavy's A Theory of Discourse (1971), Aims and Audiences in Writing (1976), and Writing--Basic Modes of Organization (1976). The treatment of the purposes and modes of discourse in the Rodi syllabus was supplemented by parallel treatments, when available, in Adelstein and Pival's The Writing Commitment (1976) and by exemplary readings from Decker's Patterns in Exposition 6 (1978). Instruction in using the conventions of Standard Written English in the context of whole discourses was provided by Corder's Handbook of Current English (1978).

Students enrolled in this option wrote during the course of the semester the equivalent of six essays (500-650 words each), not including the library paper, the journal, or the two pretest essays and two posttest essays required by the research design (see below). Of the six essays written to meet the course requirement, generally two or three were written in class and the remainder were written outside of class. Students enrolled in this option had but one teacher and met regularly in classes of 25 students each throughout the Fall Semester, some on a three-day-per-week schedule with 50-minute class periods and some on a two-day-per-week schedule with 75-minute class periods.

The Synthetic Option. The second major E 306 option, the option serving only about 3.5% of the students enrolled in E 306 during the Fall of 1978, employed a meristic approach to teaching composition. This option, rather than working from the whole discourse to its individual parts as in the analytic option, approached the development of writing skills in a more deliberately synthetic way, working from individual discourse parts--in particular, sentences--to paragraphs and, finally, to multiple-paragraph essays. The rhetorical theory underlying this second E 306 option was Francis Christensen's generative rhetoric developed through the six essays included in his Notes Toward a New Rhetoric (1967), adapted by Grady (1971) to the whole essay, and taught with further adaptions through Wittig's Steps to Structure (1975), the major textbook in the course.

The basic idea underlying this approach to composition is that of the addition of increasingly more specific details, comparisons, etc. to the base clauses of sentences, to the topic sentences of paragraphs, and to the thesis statements of essays. According to this underlying rhetorical theory, the sentence is a structural and developmental microcosm of the essay. Applied pedagogically in the second E 306 option, the theory meant that by teaching sentence structure and development first, one would, in fact, be teaching a miniature version of paragraph and essay structure and development. Accordingly, skills developed in writing sentences could be synthesized into the larger, more complex skills required, first, in writing paragraphs and, second, in writing essays. Because the text materials for the synthetic option rely on considerable grammatical terminology, teachers in this option used Blumenthal's English 3200 (1972), a programmed gramm text. Initial instruction in English 3200 was intended to provide students with the competencies that would enable them to complete more sophisticated sentence-level exercises available through an interactive computer program developed under the direction of Wittig. These exercises, in turn, prepared students to handle the treatments of paragraphs and essays contained in Wittig's Steps to Structure. With this synthetic approach to the development of writing skills, the students enrolled in this second E 306 option did not write nearly so many essays--in fact, only two besides the pretest and posttest essays collected specifically for the comparative evaluation.

The students in the synthetic option did, however, produce, through the large number of sentence and paragraph exercises, a substantial number of words during the course of the semester.

Not only did this second E 306 option differ from the first along the dimensions of underlying rhetorical theory, but it also differed with respect to instructional media. Whereas the instructional medium of the analytic option consisted of conventional instruction offered in regularly scheduled lecture-discussion classes, the instructional media of the synthetic option consisted of individualized instruction offered with tutors in a laboratory employing computer-assisted instruction. Thus while one teacher in an analytic course worked with 25 students, one teacher and three tutors, who were UT English graduate students without the M.A., worked with the 25 students enrolled in each of the sections of the synthetic option. The effect of these differences in instructional media was, of course, to change the student-teacher ratio at the beginning of the semester from 25-to-1 in the analytic classes to about 6-to-1 in the laboratory classes. Because of the attrition rate in the synthetic courses, this latter ratio fell to about 4 1/2-to-1 by the end of the semester.

1.3 Use of Data Collected During the Evaluation

Although the present study was designed specifically to compare the relative effectiveness of these two E 306 options, the data collected also enabled the investigators to study other research questions. In the present report, those data are used to develop a "profile" of E 306 students, based partly on features of "high" and "low" quality E 306 essays and partly on measurements of E 306 students in general. Those data are also used to identify a number of predictors and correlates of writing quality. The analyses of these data should prove useful in examining, for example, freshmen English exemption procedures based on standardized test scores. They should also prove useful to those at the University of Texas charged with the responsibility of teaching writing and making administrative decisions regarding the teaching of writing.

CHAPTER 2

RESEARCH DESIGN AND PROCEDURES

2.1 The Design of Ten Comparisons

In all, approximately 500 E 306 students were involved in the study reported here. This large number of students was required because the two E 306 options we wished to compare were so remarkably different. As was pointed out above, the two courses differed with respect to underlying rhetorical theory, instructional media, and textbooks or text materials. The tree diagram in Figure 2.1 represents these differences graphically by showing the number of courses which would have had to be designed and offered simultaneously in order to control for all the major variables involved in the two E 306 options. The two options themselves are pictured, respectively, in the lower left right of the tree diagram and marked with asterisks. Figure 2.1 also suggests something of the logistical problems which would obtain should such a controlled comparison be attempted systematically: had we attempted to test for the effect on student

Insert Figure 2.1 about here.

performance of all instructional variables, some 42 distinct E 306 options would have had to be offered. Obviously, such was not possible. We did, however, believe that we should, in addition to testing for the overall effect of the two E 306 options in question, test for the effect of the differing curricula in use in the two options and for the effect of the widely divergent instructional media. In addition, it was determined that we ought to study the effect on student performance within the analytic option of the Rodi syllabus.

The Five Principal Comparisons. The research design thus called for five specific comparisons, those identified in Figure 2.2, which extracts the relevant courses from the comprehensive tree diagram pictured in Figure 2.1.

Insert Figure 2.2 about here.

As Figure 2.2 and the appropriate legend show, comparison #1 tests for the relative effectiveness of the two E 306 options. Comparison #2 tests for the effect of "course content," focusing on the analytic curriculum as taught through the Rodi syllabus and the synthetic curriculum as taught through Wittig's Steps to Structure. In comparison #2, it will be observed, the instructional medium is controlled, in this case conventional classroom instruction. Comparison #5 is identical to comparison #2, except that in this case the instructional medium differs.

In comparison #5, the instructional medium is tutorial rather than conventional. Comparison #4 tests for the effect of instructional medium on student performance in the analytic course, with half of the students enrolled in a tutorial course and half in a conventional course; and comparison #3 tests for the effect of the Rodi syllabus on student performance in the analytic option. In this last comparison half the students used The Writing Commitment in a course guided by the Rodi syllabus, and half the students did not use The Writing Commitment but were exposed to the Rodi syllabus.

Sample Population and Control of Teacher Variable in the Five Principal Comparisons. Each of the five principal comparisons involved four sections of E 306 students, with approximately 50 students, or two sections, undergoing one instructional treatment and about 50 students, or two sections, undergoing the other treatment. Thus each of the five comparisons involved approximately 100 students. The students were not specifically selected at random from the population of E 306 students. However, of the 20 sections involved in the evaluation, 14 were randomly selected. The remaining six sections were made up of students who were, for the most part, channeled into the synthetic option by their advisors, who apparently believed that certain students--many of them athletes--by virtue of poor academic records or because of job-commitments would find the laboratory-tutorial course more suitable to their needs. With the possible exception of three or four of the six sections just noted, we are relatively confident that the sections involved in the study are representative of the larger population of UT E 306 students. In fact, when the pre-enrollment scores on the Scholastic Aptitude Test, the English Composition Test, and the Test of Standard Written English of the 500 students selected for the present study and of all E 306 students were compared, no statistically significant differences were found. In each of the five principal comparisons, the teacher variable was controlled. Each comparison involved two teachers, and each teacher in a given comparison taught one section on each side of the comparison. Thus altogether, 10 teachers, not including the tutors who worked in the laboratory sections, and approximately 500 students were involved in the study. These ten teachers all had previous experience teaching freshman writing, averaging 8 semesters in the college classroom, and all either had a Ph.D. or were nearing the completion of one.

The Five Derivative Comparisons. By pooling data from sections in which the same variable was controlled and by eliminating the controls exerted on the teacher variable in the five principal comparisons, we were able to create five derivative comparisons. Some of these derivative comparisons allowed us to address questions which could not have been addressed by means of the principal comparisons specified in the original research design. Other comparisons, following the removal of controls on the teacher variable, merely allowed a larger number of subjects to be used in the analysis, thus allowing us to answer certain questions using tests which were statistically more powerful. These five derivative comparisons are the following: (1) the analytic curriculum taught in a conventional classroom setting compared with the synthetic curriculum taught in a conventional classroom; (2) the analytic curriculum taught in a conventional classroom setting compared with the synthetic curriculum taught in a tutorial setting; (3) conventional classroom instruction, regardless of underlying rhetorical theory, compared with tutorial instruction, regardless of underlying rhetorical theory; (4) the analytic curriculum, regardless of instructional media, compared with the synthetic curriculum, regardless of instructional media; and (5) the analytic curriculum taught in a conventional classroom setting, compared with the analytic curriculum taught in a tutorial setting.

2.2 Pretest and Posttest Data

The research design for the present study was a pretest-posttest design. However--and this is a major weakness in this study as it is in nearly all studies of composition curricula to date--no control groups as such (see Campbell and Stanley, 1966) were provided for. That is to say, all students for whom data was either available or from whom data was collected were, during the Fall Semester of 1978, undergoing composition instruction of one of the kinds described above. Although we would have preferred to have had at our disposal a control group which was not undergoing some instruction in composition and at the same time was experiencing all of the other aspects of college life, university regulations and enrollment practices precluded our identifying and studying such a group. Hence when gains across time are reported for the various instructional treatments studied or when one instructional treatment is found to be more efficacious than another, it is important to keep in mind that we simply do not know whether a group of comparable students undergoing no instruction in composition would have performed as well or worse than the group or groups in question at the end of the semester.

Measures. Pretest and posttest essays were analyzed both qualitatively and quantitatively. Essays were rated for overall quality by experienced teachers of writing in a holistic evaluation described later in this chapter. The The essays were also submitted to quantitative analyses. The number of words in each essay was counted, and four syntactic measures used in previous research studies of college writing were computed. These four measures are described in Chapter 3. Besides the essay measures, pretest and posttest data were collected on three subtests of the *McGraw-Hill Writing Test*--language mechanics, sentence patterns, and paragraph patterns--as well as the total score on the *Writing Test*. Data were also collected for the paragraph comprehension subtest of the *McGraw-Hill Reading Test*. These data allowed us to measure the development of writing-related skills, especially usage and mechanics, and to gauge the effect of E 306 on reading skills. Writing anxiety, a trait that has been shown to affect writing performance, was measured by the *Daly-Miller Writing Apprehension Test*.

The Writing Topics. The instruments used to collect the pretest and posttest ess data were selected so that performance could be measured along a number of dimensions--(1) writing performance on a "personal experience" writing assignment which depended heavily on narration and description; (2) writing performance on an "argumentative" writing assignment; (3) mastery of basic writing skills as measured on an objective test; (4) mastery of reading comprehension skills as measured on an objective test; and (5) writing apprehension or anxiety as measured on an objective instrument. The first four and last four class meetings of a MWF class were devoted to collecting the data analyzed subsequently in this report: two pretest essays and two posttest essays; the McGraw-Hill Writing Test, Form A and Form B (Raygor, 1970b); the McGraw-Hill Reading Test (paragraph comprehension section only), Form A and Form B (Raygor, 1970a); the Daly-Miller Writing Apprehension Test (WAT) (Daly and Miller, 1975), pre and post.

The first set of writing topics, the A_1/B_1 Topics, were those used as pretest-posttest measures in the Miami University sentence-combining experiment (Morenberg, et al., 1978). They read as follows:

Topic A_1

Each of us behaves differently when we move from one group to another. We play different roles in different situations. For example, we do not act at home precisely as we act on dates, in the classroom, or before an employer. Nor do we behave with a single friend as we behave with a group of friends.

In an extensive and detailed essay, develop your ideas about the changes in our behavior. Use specific illustrations from your personal experience, from observations of others, or from books, movies, and television. You may want to explore questions of your own or answer questions like these: Why do we act differently in different situations? Are the changes in our behavior motivated by some need? Are other people misled by our behavior changes? Do such changes indicate something insincere or hypocritical about us? What happens when we do not change our behavior from one situation to another?

Topic B_1

Each of us likes to escape from reality. We often make believe that our world is different than it is. For example, we imagine ourselves reigning as homecoming queen or serving as class president, scoring the winning touchdown, or dating the most popular person in school. We identify with heroes and heroines of adventure movies, romantic stories, science fiction, and athletic events.

In an extensive and detailed essay, develop your thoughts about escaping from reality. Use specific illustrations from your personal experience, from observations of others, or from books, movies, and television. You may want to explore questions of your own or answer questions like these: Why do we escape? Is our desire to do so motivated by some need? Do you think that escape is irresponsible? harmful to ourselves and society? beneficial? necessary? satisfying? Do we escape in different ways? Do we escape less often and less completely as we mature? What happens if we are unable to escape?

The second set of topics were designed by Cameron and Witte to ellicit written discourse of a different kind. With the A_2/B_2 Topics, it was hoped that the E 306 students would adopt a position on a controversial topic, defend that position, and try to urge an audience to adopt a similar position. These topics read as follows:

Topic A_2

One issue commanding a great deal of attention nationally is the significant decline in the ability of high school graduates to read and to write. At least one solution to this "literacy crisis" has been proposed. The proposed solution would require high school students to demonstrate that they can read and write well before they graduate from high school. As might be expected, reactions to this proposed solution are divided. Some people insist that such a requirement would be unfair to the students because knowing how to read and write well is not necessary in many of the jobs for which high school graduates qualify. Other people insist that employment skills are not at issue. These other people insist that by knowing how to read and write well the individual is more likely to continue growing intellectually, socially, and culturally after his or her high school years.

In a multi-level essay (that is, an extensive and detailed essay), argue objectively for one point of view or the other. Your argument should be well organized and easy to follow. It should be unified and complete, and it should be free of all logical fallacies or errors in reasoning. Support for your position should come from your knowledge of the issue and its implications, and that support should be presented objectively without the use of first-person pronouns. Before you begin to write, you may want to explore questions of your own or to explore questions such as these: Are reading and writing essential to the full development of the individual? Would an individual deficient in reading skills be more likely to be misled by television commercials, by newspaper and magazine advertisements? How do well-developed, or underdeveloped, reading and writing skills affect people's lives? Is there any relationship between reading and writing skills and the ability to appreciate cultural artifacts such as movies, poems, popular songs, novels, or television shows? Are reading and writing related to employment opportunities? to successful employment?

Topic B_2

Composition courses are required in many high schools for most students. Many people, both educators and non-educators, believe that composition courses play an extremely important part in the educational process. These people argue that composition courses develop writing skills which will serve the student well both in other courses and in the world of work. However, in the last few years, composition courses have been severely criticized because, the critics assert, they do a poor job of teaching students to write. Some people insist that the problem is one of relaxed standards in high school composition courses. These people also insist that the effectiveness of high school composition courses could be increased if the courses made greater demands on the student, thus making it more difficult for the student to earn passing marks or grades. Other people assert that composition courses in high schools are not related to the world outside the classroom. According to these people, composition courses should be abolished because they have no practical value.

In a multi-level essay (that is, an extensive and detailed essay), argue objectively for one of the above three points of view. Your argument should be well organized and easy to follow. It should be unified and complete, and it should be free of all logical fallacies or errors in reasoning. Support for your position should come from your knowledge of the problem and its implications, and that support should be presented objectively without the use of first-person pronouns. Before you begin to write, you may want to explore questions of your own or to explore questions such as these: Is there a relationship between what the student does in a high school composition course and the world outside the classroom? Do high school composition courses have a value apart from the teaching of writing? What standards should be used in assessing student progress in high school composition courses? In what ways are composition courses practical? impractical? What should the goals of a high school composition course be? Why are some high school composition courses effective? ineffective? Is writing related to other language skills? What role should composition courses have in a high school curriculum? Is knowing how to write well vital to the future employment of high school graduates? vital to the future employment of some but not of others?

Administration of the Test Instruments. These sets of writing topics were used within the pretest-posttest design. Although it was believed that the two topics in either set were roughly comparable with respect to levels of difficulty, it could not be determined in advance whether administering one topic in either set as a pretest topic and the other as a posttest topic would bias the results. If one topic in a set were more difficult for the students to write on than the other and if that topic were administered only as a pretest, then the resultant change scores across time would likely have been inflated. Such inflation could suggest that the variable being tested in a given comparison had a more powerful effect on student performance than it actually did. Conversely, if the two topics in a given set were not of comparable difficulty and the easier topic were administered only as a pretest topic and the more difficult one only as a posttest topic, then those resultant change scores could also have been misleading. In that event, the results may have indicated that no positive change had occurred as a function of the instructional treatment or the curricular design when, in fact, the absence of positive change would have been an artifact of administering the more difficult topic as a posttest topic only.

To guard against these and other possible sources of bias, we administered one topic in a set as a pretest to one section in each half of each of the five principal comparisons and the other topic as a pretest to the other section. For the posttest essays, the process was simply reversed. Table 2.1 shows the way the topics were administered within each of the five principal comparisons. The section numbers are those assigned by the investigators, and the brief descriptions are provided to help distinguish between the sections within each of the five principal comparisons. This same procedure was followed in the administration of the McGraw-Hill tests, with those students writing on the A_1 topic as a pretest completing Form A of each test and those writing on topic B_1 completing Form B of each test as a pretest. For the posttest the forms were reversed. Such stringent controls against one writing topic within a given pair of topics, or one form of a McGraw-Hill test, biasing the derivative comparisons could not, of course, be maintained for the derivative comparisons. The various sections of E 306 involved in the derivative comparisons are specified in Table 2.2; and if Table 2.2 and Table 2.1 are viewed together, this lack of control over the writing topics and the forms of the McGraw-Hill tests as far as the derivative

---

Insert Table 2.2 about here.

---

comparisons are concerned can be observed. It is also important to reiterate that while the teacher variable was rigorously controlled

Insert Table 2.2 about here.

in the five principal comparisons, such control was not possible in the five derivative comparisons.

Sample Selection. Although pretest and posttest essays for each set of topics were collected in the manner indicated by Table 2.1 for each of the approximately 25 students in each of the 20 sections, extremely limited financial resources precluded evaluating all of the 2,000 essays (four essays each from each of 500 students) collected. Thus before holistic evaluations of the essays took place and before syntactic analyses were performed on them, a random selection of nine students from each of the 20 sections was made following methods outlined by Plutchik (1974) and others for the generation of random numbers from a finite set. Although complete sets of essay data were available for these 180 students, complete sets of objective test data were not available for these students. Complete pretest and posttest data for the McGraw-Hill Writing Test were not available for 15 of the 180 students; and complete pretest and posttest data for the McGraw-Hill Reading Test and the Daly Miller Writing Apprehension Test were not available for

21 of the 180 students. These missing data sets were spread across the 20 E 306 sections used in the present study.

2.3 Holistic Evaluation Procedures

Preparing the Essays for Holistic Evaluation. Prior to submitting the 720 essays to holistic evaluation, all identifying marks were either removed from or blacked out on the essays. Such marks included the student's name, the date, the E 306 "unique" number, the teacher's comments, and the teacher's grade. This procedure was necessary to insure the anonymity of the individual students and to avoid biasing the raters who would score the essays. This process was, in many cases, a very time-consuming one because many of the papers had been heavily annotated by the respective teachers. Each of the 720 essays was then coded on each page with a six-digit number, the meaning of which was known only to the investigators. This code number allowed the investigators to identify the section from which a given essay had been selected, the student within the section who had written the essay, the time of testing (whether pretest or posttest), and the essay topic. Thus the code number "120664" would indicate to the investigators that the essay come from section 12 (which was itself a code for a give "unique" number), that it was written by student "6" in section 12, that it was a posttest essay (if fifth digit in the code number had been a 5, the essay would have been a pretest essay), and that it was written on the B_2 topic (as the sixth digit, 1 = A_1, 2 = B_1, 3 = A_2, and 4 = B_2). After all identifying marks had either been removed from or blacked out on each essay and after all 720 essays had been coded as indicated above, the essays were photocopied before being submitted to single-impression holistic evaluations.

Single-Impression Holistic Evaluation. This type of holistic evaluation is a method developed by the Educational Testing Service (ETS) for assessing large numbers of students writing samples relatively quickly and has been proffered as the most reliable and least costly method developed to date for assessing student writing. In single-impression holistic evaluation, pretest and posttest essays written by the same group of subjects on like or comparable writing topics are pooled. Raters are not informed which essays were written as pretest measures and which were written as posttest measures. Holistic evaluations are preceded by intensive training sessions. These are probably an integral part of holistic evaluations. During these sessions, raters read essays written on the same topics as those to be evaluated, rate the essays, and then compare ratings. More papers are read and rated, and the ratings compared among raters. This process continues until all raters are rating virtually every paper the same way. The process itself palces a great deal of pressure on dissenting raters, those whose ratings depart markedly form the majority rating assigned to a particular paper. Typically, holistic rating sessions employ either a four-point, a six-point, or an eight-point scale, with the number "1" assigned to the essays of the lowest quality and the highest number assigned to the essays of the highest quality. Because the larger the scale the more difficult it becomes to achieve acceptable inter-rater reliabilities (see below), we employed a four-point holistic scale in the present study.

In holistic evaluation, raters are forced to read essays very quickly and to make impressionistic judgments as to the relative quality of an essay, presumably not allowing particular features to be weighted more heavily in their decision-making processes than other features. For example, raters are discouraged from assigning a low score to an essay simply because the writer did not know, say, how to avoid comma splices or how to make pronouns agree with their antecedents in number. Rather, the emphasis is always on the "total impression," the holistic impression, of the essay. Raters are also instructed and trained to ignore handwriting and spelling errors, a problem which often takes some time to overcome in training sessions.

Another difficulty encountered in holistic training sessions and in holistic rating sessions (if the training sessions are inadequate) is the tendency of raters to equate holistic scores with grades. Holistic scores differ from grades in at least three important ways. First, grading can be described as applying criteria that have been abstracted from some standard of performance not derived from the particular essays being graded but rather imposed upon those essays. These external criteria have, of course, been shown to differ from one grader to the next. If holistic raters were allowed to treat holistic scores as though they were grades, each rater would probably bring a different set of internalized and hierarchial criteria to the evaluation. In holistic evaluation, criteria must be generated from or identified in the referent group being evaluated in order to achieve agreement among raters. That is to say, during the training sessions raters must be exposed to the range in quality of the set of papers to be submitted to holistic evaluation. From these representative papers (which in the present study were drawn from the 1,280 papers not selected for evaluation), the raters extract the criteria which apply to that set of essays. Thus holistic scores are always made to be relative to only those papers being evaluated. Second, grading techniques are seldom, if ever, applied to a set of essays consisting of essays separated in time by several months. Third, in holistic evaluation, where pretest and posttest essays are pooled, scores should be distributed across each scoring category, something which seldom occurs in classroom practice. By attempting to force essays into each available category, it becomes possible to achieve the spread of scores required to show that X_1 number of essays were better than X_2 number of essays. When a given holistic session is completed, approximately 50% of the pooled pretest and posttest essays should fall above the mean and 50% should fall below the mean. If a given instructional treatment had any positive effect on student writing performance, more posttest essay scores should appear above the mean for that group of students than below.

The holistic rating sessions employed in the study being reported here consisted of six to ten raters, depending on the particular rating session. The photocopied and coded essays for each set of essay topics were pooled and randomly sequenced with regard to E 306 section, time of writing (pretest or posttest), and topic within essay set. The essays were then divided into stacks which totalled twice the number of raters involved in a given rating session. Each stack was then read and scored by two raters. Neither rater reading a given stack of essays had knowledge of the scores assigned those essays by the other rater. Because a four-point scale was employed with two raters for each essay, the summed scores for a given essay ranged from two to eight. When the two scores assigned a given essay differed by two or more on the four-point scale, the essay was submitted to a third reading by a third rater who had no knowledge of the scores assigned by the previous two raters. The third rater's score was then added to the previous rating to which the third score came closest. In cases where a two-four split or a comparable split appeared after the first two ratings and the third rating fell between the two earlier scores, the scores more distant from the middle were used. For example, if an essay had been rated "1" and "3" by the first two readers and "2" by the third reader, the "1" and the "2" were summed. Interrater reliabilities were computed to determine with what consistency the raters assigned holistic scores to the essays. Using a technique described by Ebel (1951), we computed separate reliability coefficients for the holistic scores assigned to the A_1B_1 pretest essays, the A_1B_1

posttest essays, the A_2B_2 pretest essays, and the A_2B_2 posttest essays. These coefficients were found to be 0.76, 0.78, 0.82, and 0.84, respectively.

The Distribution of Holistic Scores by Scoring Category. As was pointed out above, a four-point scale was employed in the holistic rating sessions and scores from two raters were used to arrive at a summed score recorded for each essay rated. The distribution of scores within the four scoring categories appears in Table 2.3 and Table 2.4. Table 2.3 reports on the actual number of scores in each category and Table 2.4 converts those data to percentages of total scores awarded. Because two such scores were used for each of the 720 essays rated, the number of scores included in Table 2.3 is double the number of essays rated.

Insert Table 2.3 about here.

Insert Table 2.4 about here.

When the scores (with a possible range of one to four) assigned a particular essay are summed, the number of scoring categories increases from four to seven, and a scoring range from two to eight results. Table 2.5 and Table 2.6 present the distribution of these summed scores assigned in the holistic rating sessions to the 720 E 306 essays. These same data, con-

---

Insert Table 2.5 about here.

---

verted into percentages of total scores in each scoring category (2 through 8), are presented in Table 2.6. As with Table 2.3

Insert Table 2.6 about here.

---

and Table 2.4, Table 2.5 and Table 2.6 give some indications of improved writing performance in the 20 E 306 sections during the Fall Semester of 1978. If we ignore the scores falling midway (that is, those scores in scoring category 5) between the extremes of scoring categories 2 and 8, we find that 45.5% of the scores on the A_1B_1 pretest essays fall into the low range (scoring categories 2, 3, and 4) and that only 22.8% of the scores fall into the high range (scoring categories 6, 7, and 8). When the A_1B_1 posttest essays are viewed in the same way, we observe that a significantly smaller percentage, 39.9% as compared to 45.5% of the essays fall into the low range and that a significantly higher percentage, 37.8% as compared to 22.8%, fall into the high range. This difference in percentages, again, indicates that the writing performances, as measured by the A_1B_1 writing topics, of some of the 180 E 306 writers studied were better at the end of the semester than at the beginning. When we examine the distribution of pretest and posttest scores on the A_2B_2 writing topics, we note a greater improvement over the course of the semester.

Of the A_2B_2 pretest scores, 51.6% fall into the low range and 26.1% fall into the high range. Of the A_2B_2 posttest scores, only 37.2% fall into the low range while 44.9% fall into the high range. If we pool the A_1B_1 pretest scores and the A_2B_2 pretest scores and compare their distribution with that of the pooled A_1B_1 posttest scores and A_2B_2 posttest scores, we get some indication of the overall change in writing performance, as measured by the two sets of essay topics, for the E 306 students involved in the present study. Of the pooled pretest scores, 48.6% fell into the low range and only 24.5% fell into the high range; but of the pooled posttest scores, only 38.1% fell into the low range while 41.3% fell into the high range. Thus for the pooled scores we see that over the course of the semester 10.5% of the E 306 students studied moved out of the low range of writing performance while 16.8% moved into the high range.

While these percentages give a very general impression of improvement in writing performance across the time of one semester for all 180 E 306 students involved in the study, our principal concern is the instructional variables which caused the most improvement or which affected writing performance adversely. The data in these four tables are presented simply to indicate the workings of holistic evaluation procedures. In subsequent chapters of this report, we will focus directly in the apparent causes for the improvement suggested by these four tables.

2.4 Syntactic Analyses of the E 306 Essays

The single-impression holistic evaluations of the 720 E 306 essays provided an indication of the quality of each essay in relation to the referent group of which that essay is a part. But holistic scores indicate very little about the specific features of essays which can and do change across time. Although we would have liked, for example, to have submitted all 720 E 306 essays to analyses of semantic and lexical features and to analyses of developmental and organizational patterns, such analyses either required much more money than was available to us or required analytic techniques much more sophisticated and efficient than those which have been developed to date. We did, however, examine certain features of all 720 essays which were submitted to holistic evaluation, partly because some of those features have been shown to predict writing quality and partly because the analyses performed would allow us to place the changes observed in the E 306 essays in the context of several previous studies.

In Chapter 3 of this report, we review the literature relevant to our analyses of the syntactic features of the 720 E 305 essays. Since most of the syntactic analyses we performed are based on the total number of words in an essay, it was necessary, first, to count all the words in each of the 720 essays which contained, on the average, 433.91 words each. The rules observed in counting the words in each essay were those observed in the Miami University Sentence-combining experiment (Morenberg, et al., 1978) and in Faigley's (1979 b,c) study of the effects of generative rhetoric on writing performance.

CHAPTER 3

SOME FEATURES OF "HIGH" AND "LOW" QUALITY E 306 ESSAYS

A question of continuing interest to researchers in writing is what internal characteristics distinguish essays ranked high and low according to overall quality. Research at the college level has for the most part taken two approaches to this question, examining errors (e.g., Shaughnessy, 1977) and syntactic features (e.g., Gebhard, 1978). Neither of these approaches has been entirely satisfactory. Maimon and Nodine's sentence-combining experiment (1978) suggests that, as is true for the acquisition of other skills, certain kinds of errors accompany certain stages in learning to write. The underlying causes of error are often complex and difficult to trace. Consequently, researchers can conclude little more than what is obvious: low-rated papers usually contain far more errors than high-rated papers. Likewise, Gebhard found that with few exceptions the syntactic features of high- and low-rated essays written by college students are not clearly differentiated. Gebhard saw no specific directions in her data for teaching the improvement of written syntax. Another approach to distinguishing essays rated high and low in quality is by content analysis. Lunsford (1979) found that essays rated low in quality typically contain more first- and second-person pronouns than high-rated papers. Lunsford's findings tend to support the more general observation that college freshmen with underdeveloped writing skills show more evidence of cognitive egocentrism than their counterparts with more highly developed writing skills (cf. Barritt & Kroll, 1978; Lunsford, 1977).

The purpose of the present chapter is to suggest some particular ways in which the better E 306 essays differ from the poorer ones. This purpose, it is hoped, will benefit those concerned with matters pertaining to curricular and instructional design for E 306 courses, providing an approximate profile of the best and the poorest incoming E 306 freshmen writers. This "profile" will be based on a number of detailed analyses --an error analysis, an analysis of certain syntactic features, an analysis of the frequency of use of first- and second-person pronouns, and an analysis of cohesive devices. Such detailed analyses are, of course, both extremely tedious and time-consuming, requiring that the number of writing samples be severely limited. Because of the necessarily limited number of writing samples examined in such detail, we can only hope that the analyses reported here are generalizable to the E 306 population at large.

From the pretest essays we selected five of the poorest essays written on the A_1 topic and five of the best written on the same topic. "Poorest" and "Best" were determined by the scores assigned the respective essays during one of the holistic evaluation sessions. The five "poorest" essays all received summed scores of 2, while the five "best" essays received summed scores of either 7 or 8. Descriptive data focusing on length factors of the two sets of essays appear in Table 3.1

Insert Table 3.1 about here.

As Table 3.1 indicates, the "high quality" essays were more than twice as long as the "low quality" essays, whether viewed in terms of average length in words, sentences, T-units (according to Hunt [1965] who introduced the term, a T-unit is a main clause plus all subordinate elements attached to or embedded in it), or clauses. We will have many occasions in the pages which follow to discuss the importance of the differences in length between the "low" and "high" quality essays, between the "poorest" and the "best" essays.

The ten essays represented in Table 3.1 were analyzed in considerable detail in order to identify some specific differences between the "best" and "poorest" essays in the sample. By identifying a range of features and by suggesting how these features relate to writing quality, we will be better able to recommend certain instructional techniques and practices which should help writers of "poor quality" essays to become more like their counterparts who write essays of much higher quality. In this analysis, we focus on the occurrence of usage "errors," on respective levels of syntactic complexity, on the frequency of first- and second-person pronouns, and on text cohesion. Each of these topics will be addressed separately in the pages which follow.

3.1 Error Analysis of E 306 Essays

The analysis of usage "errors" in student writing is important for several reasons. First, usage "errors" or "mechanics", as they are often called, are frequently cited by educators outside English departments, by the business community, and by the public at large as evidence of a decline in the writing abilities of college students and graduates alike. Second, teaching students to write using the dialect of standard Written English appears to be one of the many functions expected of service-oriented composition courses and programs. Third, several members of the English faculty at the University of Texas have cited "mechanics" as a major problem among English 306 students. Fourth, problems in usage or mechanics frequently point to writing problems of a more severe nature.

Following procedures comparable to those used recently in a study of the writing of entering freshmen at SUNY-Buffalo (Cooper, et al., 1979), we concentrated on three categories of "error," addressed below under the headings of Punctuation, Grammar, and Spelling. Before presenting the results of our analyses of each class of errors, we will indicate the parameters of these classes, providing examples of each type of error in each class from the essays we analyzed.

Punctuation. Punctuation errors were broken down into two subclasses, one including the punctuation expected at the ends of sentences, and one including punctuation within sentence boundaries.

Errors in sentence end-stop punctuation were typically of two kinds, the comma splice and the fragment.

The comma splice occurs when a writer joins two independent clauses with a comma, omitting the necessary coordinating conjunction. The following group of words contains a comma splice.

(1) Some behavior is repressed, not admitted to the consciousness, often this causes mental problems.

According to composition handbooks, the comma appearing between "consciousness" and "often" is not sufficient to join the preceding and following independent clauses. If the sentence were punctuated "correctly," "often" would need to follow the coordinating conjunction and or a semi-colon which would function as a period.

The fragment, on the other hand, fails to meet one or more of the minimum requirements of sentences. Typically, a fragment either (1) lacks a finite verb in the predicate position or (2) is preceded by a subordinating conjunction such as although, when, because, or if.

Consider the following two series of words, both punctuated as sentences by E 306 students.

(2) Not because of the fact that he is, in reality, that kind of person, but because that is the way the situation dictates it.

(3) The main one being the need to be accepted.

In example 2, both of the long clauses are governed by the subordinating conjunction "because," thus making them subordinate to some independent clause not included between the initial capital letter and the terminal period. In example 3, "being" is not a finite verb, thus leaving the subject, "one," of the series without a grammatically acceptable predicate.

In addition to the comma splice and the fragment which result from "errors" in punctuating sentence boundaries, other major punctuation errors occur within sentences themselves. Although writers have many opportunities for incorrectly punctuating within sentences, among E 306 students a few types of "within sentence" punctuation errors predominate. These errors may result from the writer either including unnecessary punctuation or omitting the necessary punctuation. A few typical examples from the student essays will illustrate.

(4) A friend of mine, for example, acts, so shy people think he is the quite [sic] type.

(5) For instance, the love, that all of us feel at one time in our lifes [sic], has the ability to bring two people very close together. . . .

(6) However it is obvious that human behavior is constantly changing. . . .

In example 4, the comma after "acts" is superfluous, serving no function whatsoever. In example 5, the relative clause, "that all of us feel at one time in our lifes," is punctuated as though it were a non-restrictive clause when, in fact, it is a restrictive clause and should thus not be set off with commas. In example 6, the writer has failed to set off with a comma the introductory conjunctive adverb, "however," as handbooks of usage direct.

Grammatical Errors. Of the many possible types of grammatical errors, six appear frequently in student writing. These errors are errors in verb tenses, errors in number agreement between a verb and its subject, errors in pronoun reference, errors in pronoun number agreement, and errors in modification in which a modifier is either misplaced or left dangling with no noun-actor for it to modify. Again, examples from the students' papers will help illustrate these types of grammatical error.

(7) Love, which is one of the rarest emotions know to us, is constantly a topic of are [sic] thoughts.

(8) Therefore, society and culture encourages this role-playing game.

(9) If you meet a person and they are really nasty to you for no apparent reason you will no doubt have harsh feelings toward that person.

(10) By allowing extremely violent television shows and movies it can be expected that our national crime rate will increase. . . .

In (7), the present tense "know" appears where the past participial form of the verb would normally appear; and in (8) a singular verb, "encourages," is incorrectly supplied for a plural subject, "society and culture." In (9), the plural pronoun "they" is used to refer to a singular entity, "person." And in (10), "By allowing extremely violent television shows and movies" stands next to an expletive "it" which cannot perform the action indicated by the verbal "allowing," thus leaving the entire introductory participial phrase "dangling" apart from an actor capable of "allowing."

Spelling Errors. Finally, we counted misspelled words. As might be expected, both the "high" quality and the "low" quality writers experienced the most difficulty with homonyms--words sounding the same but having different meanings--such as our-are, there-their, their-they're, there-they're, and it's-its. Two other particularly troublesome sets, though not strictly homonymous, were then-than and affect-effect. In arriving at the number of misspelled words in each paper we counted a misspelled word only once in any given paper, even though that word might have been misspelled several times by the writer.

Comparison of "High" and "Low" Quality Essays. Because there was considerable variation in the lengths of the essays in both the "high" quality set and the "low" quality set, it was impossible to make any comparisons between the two sets without some common denominators. We chose two common denominators, words and T-units, for each set of five essays. By dividing the number of errors of each type for a given set of essays into the total number of words of the essays in that set, we were able to calculate the frequency of occurrence, in words, for each error type. Similarly, by dividing the number of errors of each type into the total number of T-units of the essays in a given set, we were able to calculate the frequency of error in terms of T-units. The results of these calculations appear in Table 3.2 and Table 3.3.

Insert Table 3.2 about here.

Insert Table 3.3 about here.

Table 3.2 and Table 3.3 present the same set of raw data (the number of usage errors) in two different ways, one in terms of the mean number of words separating errors of various types and one in terms of the mean number of T-units separating the errors of various types. If we look at the frequency of occurrence of all errors in terms of words and T-units, we see that the "low quality" essays exhibited an error of some type nearly three times as often--one every 29.39 words as opposed to one every 87.24 words or one every 2.15 T-units as opposed to one every 5.7 T-units--as the "high quality" essays. Some types of error, of course, appeared more often than others. For example, an error in end-stop punctuation, resulting in either a comma splice or a fragment, occurred nearly eight times as often in the "low quality" essays as in the "high quality" essays. With regard to this particular type of error, it is important to note that the fragment occurred more often in the "high quality" essays than it did in the "low quality" essays. However, in every case of a fragment in a "high quality" essay, the fragment was the result of the writer creating a long terminal sentence modifier but failing to attach it to the preceding main clause with a comma. It seems to us that this type of punctuation error could result from a writer attempting to manage a complex syntactic structure which the writers of the "low quality" essays seldom attempted. This alone should suggest that any approach to composition which stresses error reduction primarily may discourage students from attempting to manipulate complex syntactic structures with which they are not altogether comfortable. The other type of punctuation error we examined, that involving either the omission of necessary punctuation or the inclusion of unnecessary punctuation, also points to a significant difference between the two groups of writers. Whereas the group of "low quality" writers simply seemed not to understand the principles of within-sentence punctuation and thereby exhibited many different such errors, most of the within-sentence punctuation errors of the "high quality" group resulted from the omission of commas after initial conjunctive adverbs, words which do not appear very often in the "low quality" essays. Among the grammatical errors, verb tense errors appear not to be a major problem for either group, nor does the occurrence of misplaced or dangling modifiers present much cause for alarm. Pronoun reference and pronoun agreement were, however, something of a problem for both groups of writers, even though a pronoun error occurred only about two and one-half times as often in the "low quality" essays as in the "high quality" essays. Yet there were some noticeable differences in the kinds of pronoun errors most frequently observed in both groups of essays. Most of the "errors" observed in the "high quality" papers were errors in pronoun number agreement. (We hesitate to label these as errors because of their frequent use in spoken English.) In the "low quality" essays, the pronoun errors included ones involving ambiguous referents as well as number agreement. If we were asked to speculate about the correlation of pronouns with ambiguous referents to writing quality, we would suggest that such errors detract from the overall clarity exhibited in a piece of writing.

It would not be appropriate to end this section on error analysis without some attempt to suggest how these data, collected from beginning E 306 students at the University of Texas, relate to data collected from groups of beginning college freshmen at other universities. Although few such error analyses were available to the investigators, the two studies which were available that do report error analyses in detail provide the bases for some interesting comparisons.

In order to make these comparisons we pooled error-data from the "high quality" and "low quality" E 306 essays in order to extrapolate error frequencies of various types for an "average" beginning E 306 student. The results of this pooling appear in Table 3.4.

Insert Table 3.4 about here.

Given these frequencies of error, what can be said about the error-tendencies of beginning E 306 students? The first point to note is that our "hypothesized" "average" E 306 student commits punctuation and spelling errors at a rate more than twice as great as the rate with which he/she makes grammatical errors. The difference between these rates of punctuation-error frequency and spelling-error frequency on the one hand and grammar-error frequency on the other hand probably lies in the linguistic difference between spelling and punctuation, on the one hand, and grammar, on the other. Spelling and punctuation are particular to written language and are learned relatively late, while grammar--"good," "standard," "substandard," or otherwise--is an inseparable part of language itself, something every person has before he/she attends his/her first kindergarten class. Punctuation and spelling, properly speaking, are writing skills, artificial adjuncts to language usage, and as such develop much more slowly and much later. Beginning E 306 students have been speaking the language for some seventeen or eighteen years, but they have been aware of the differences between oral and written discourse for fewer than half of those years and have perhaps never been required to write regularly at any time during their lives. By the time one reaches the age of eighteen, he/she has in all probability already produced oral discourse in greater quantity than he/she will ever produce written discourse.

These differences in error frequencies or rates are, of course, also reflected in the ratio of error types to the total number of errors observed in the sample, ratios which can be expressed as percentages of the total errors. Of all the errors observed, 41% were punctuation errors and 39.8% were spelling errors. Thus nearly 81% of all errors committed in the written discourse sample were errors that would not have figured at all in the oral discourse of these same students, thus again suggesting that these E 306 students had a basically good command of important matters of usage.

But how do these E 306 students compare with other groups of beginning college freshmen? Witte and Sodowsky (1977) did a similar analysis of the written discourse produced by beginning college freshmen at Oklahoma State University. For this population, which in 1976 had mean ACT English, Mathematics, and Composite scores of 18.3, 21.3, and 20.2, respectively, Witte and

Sodowsky reported that punctuation errors of all types occurred on the average once every 91.8 words and grammar errors of all types once every 73 words. Compared to the E 306 freshmen in the present study, the OSU students were considerably less proficient in matters of usage, committing punctuation errors with a 31% greater frequency than the UT students and grammatical errors with a frequency approximately 75% greater than that observed in the writing samples collected from the E 306 students. It should, however, be noted that of the OSU population about 18% were "developmental" or "remedial" students, students who would be classified as "provisionals" at the University of Texas. There were no "provisional" students in the University of Texas population we studied.

In a study of SUNY-Buffalo beginning freshman writers (who, incidentally, were drawn from the top 12% of high school graduating classes), Cooper, et al. (1980) report on an error analysis of two sub-groups, those writing "high quality" essays and those writing "low quality" essays as determined by a holistic evaluation similar to the ones employed in the present study. To draw comparisons between that group and the E 306 student writers, we summed the results for the two SUNY-Buffalo sub-groups and divided by two. Because the results for the SUNY-Buffalo analysis are reported in terms of errors per 1000 words, we converted the E 306 data to that form, as shown in Table 3.4. As Table 3.4 indicates, for the E 306 students 7.42 punctuation errors occurred for every 1000 words, or about three punctuation errors per "average" freshman essay.

For the SUNY-Buffalo students, 9 punctuation errors occurred for each 1000 words, or 21% more often than for E 306 students. A similar difference between the two groups is observed when the respective occurrences of grammatical errors are compared. The SUNY-Buffalo writers committed, on the average, 4.5 grammatical errors for every 1000 words, while the E 306 students committed 3.49 grammatical errors per 1000 words, a frequency 29% smaller than that of the SUNY-Buffalo writers. So, too, with spelling: the SUNY-Buffalo freshmen misspelled about 9.5 words per 1000 written, while the E 306 students misspelled only about 7.2 words for each 1000 produced. When all three categories of error are collapsed for each group, we find that the SUNY-Buffalo freshmen committed about 23 errors for each 1000-word span of text written, while the E 306 students committed 18.12 errors for a text of the same length. This higher frequency of error among SUNY-Buffalo students may, however, be attributed to the fact that they tended to write more complex syntactic structures, a matter discussed in some detail in the following section of this report. Nevertheless, the E 306 writers, half of whom were judged to be "high quality" writers and half of whom were judged to be "low quality" writers, produced approximately 27% fewer errors per 1000 words of text than did the beginning college freshmen at SUNY-Buffalo.

Although the error analyses for these groups of beginning college writers--the E 306 students, the SUNY-Buffalo students, and the Oklahoma State University students--are not strictly comparable because the writing samples were collected under differing conditions on different writing topics and because the procedures for collecting the writing samples and for counting the errors may have differed in some ways not reported, it is reasonable to infer that the E 306 students at the University of Texas whose writing we examined hold their own quite well against beginning college freshmen at other universities. And this speaks well of the University of Texas freshman class at large because, it must be remembered, nearly 30% of each freshman class at the University of Texas is given advance standing credit for E 306.

3.2 Syntactic Analysis of E 306 Essays

As Faigley (1979d; in press) and Williams (1979) have argued, "syntactic complexity" (or "maturity" or "fluency") is in some sense a name in search of a concept, primarily because it is impossible to say with much confidence what specifically the indices most frequently used to calculate the phenomenon measure. It would appear that to a great extent and to many researchers "syntactic complexity" is whatever the indices used to measure it measure, a circular and operational definition not particularly useful in empirical research, if empirical research is ever to advance theory. Yet, on a very general level, "syntactic complexity" seems somehow to refer to the use or manipulation of syntactic structures in such a way that a writer is able to say more in fewer words. Thus in a very real sense "syntactic complexity" presupposes some notion of semantic density in written discourse. However, the relationship between the "syntactic" and the "semantic" has not been explored or applied in empirical research, thus making many of the results of empirical research in this area of limited value to other researchers or to teachers of writing.

While some researchers have focused their criticism on the meaning of "syntactic complexity," others have raised issues related to the indices most frequently employed to measure it. For example, Brereton, et al. (1979), have raised questions focusing on the cross-sample reliabilities of varying counting procedures, thereby suggesting that results obtained in one study may not be usefully compared to results obtained in another. Raising a different issue, O'Donnell (1976) has pointed out that "there are no data" to suggest how consistently the measures or indices most commonly employed in empirical research measure the writing of individuals in different situations. Witte, together with two colleagues, has attempted to explore the implications of O'Donnell's observation and addressed, both in terms of groups and individuals, the question of the stability of these indices within and across different modes of written discourse (see Witte and Davis, 1978; Witte, 1979; Witte and Davis, 1980 and in press). To some extent, the research of Witte and the two Davis's focuses on questions implicit in a large number of previous studies--e.g., Seegars (1933), Frogner (1933), Anderson (1960), Marckworth and Bell (1967), Johnson (1967), Veal and Tillman (1971), San Jose (1972, 1974, 1978), Perron (1976a, 1976b, 1976c, 1977), and Crowhurst (1978, 1979)--all of which point to considerable syntactic variation across discourse of different kinds.

These questions and issues notwithstanding, O'Donnell (1976) and (Hunt, 1964, 1970, 1977) present evidence suggesting that some syntactic indices can be used to track gross normative changes in syntactic manipulation for various populations. Whatever their strengths and weaknesses, different indices of syntactic complexity have been used regularly to study the oral and written discourse of students since LaBrant (1933) published her study.

Most recently, a large number of researchers have adopted Hunt's (1965) "syntactic maturity" model and its corresponding set of indices. Hunt himself posited the use of those indices as developmental or normative measures. However, the model Hunt espouses--with its range of "complexity" spanning from the writing of elementary children to that of professional writers who had published in Harper's and Atlantic--has been interpreted to mean that some kind of equivalence exists between syntactic maturity or complexity, defined in terms of Hunt's indices, and writing quality. Such is not an unreasonable interpretation, given Hunt's emphasis (1965, 1970) on paraliterary discourse at one end of his "developmental" continuum. Hunt's model and its continuum have recently been challenged on theoretical grounds by Kinneavy (1979) and on empirical grounds by Nold and Freedman (1977), Witte (1978), Marzano (1976) and Faigley (1979c).

Hunt's model of syntactic maturity or complexity is based on analyses of written discourse, analyses which use five measures or indices. Hunt and others--e.g., O'Donnell, et al.

(1967), Witte and Sodowsky (1978), Sodowsky and Witte, (1978), Stewart (1978a), and Cooper, et al. (1979)--have used these as developmental indices of syntactic complexity. Others--e.g., Mellon (1969), O'Hare (1973), Combs (1976), Ney (1976), Mulder, et al. (1978), Daiker, Kerek, and Morenberg (1978; Merenberg, Daiker, and Kerek, 1978), Faigley (1979b, 1979c), and Stewart (1978b)--have used one or more of these five measures in pedagogical research to measure the effects of certain types of writing instruction on the writing behaviors of students from the elementary to the college level.

Although there remain several troubling issues and many unanswered questions in the research on the development of syntactic complexity or maturity, it will be useful to compare our "high" and "low" quality E 306 writers to writers of different ages studied by other researchers. In this comparison, we will focus attention on gross measurements of syntactic manipulation--mean T-unit length, mean clause length, mean number of clauses per T-unit, mean number of T-units per sentence, and mean sentence length--and on the occurrence of unrestricted sentence modifiers in the terminal and the medial positions in the sentence. The gross measures are those posited by Hunt, while the others are based on the theoretical or speculative work of the late Francis Christensen (1967) and the empirical work of Nold and Freedman (1977) and Faigley (1979a, 1979c).

Mean T-unit Length. Hunt (1965) introduced the T-unit, or "minimal terminable unit," as a measure of syntactic maturity.

A T-unit is one independent (or main or base) clause plus all the subordinate elements attached to or embedded in it. According to Hunt, such a measure or index is needed because sentence length varies so greatly within and across writers punctuation skill or writing style. The following according to passage, punctuated as a single sentence by one of the elementary students whose writing Hunt (1965, p. 20) studied, will illustate the unreliability of mean sentence length as a measure of syntactic maturity or complexity and will introduce Hunt's alternative index, mean T-unit length.

> I like the movie we saw about Moby Dick the white whale the captain said if you can kill the white whale Moby Dick I will give this gold to the one that can do it and it is worth sixteen dollars they tried and tried but while they were trying they killed a whale and used the oil for lamps they almost caught the white whale.

Obviously, this is not a sentence that any teacher of college writing would like to receive from a student. But it does illustrate the unreliability of sentence length as an indicator of maturity or complexity in syntactic manipulation. Even if the several run-on, or fused, sentences were separated with semi-colons, thereby making the sentence conform better to acceptable punctuation practice, the problem remains. But if the passage were segmented into T-units, or "minimal terminable units," we would begin to get some indication of the syntactic maturity or complexity exhibited by it.

> I like the movie we saw about Moby Dick the white whale // the captain said if you can kill the white whale Moby Dick I will give this gold to the one that can do it // and it is worth sixteen dollars // they tried and tried // but while they were trying they killed a whale and used the oil for lamps // they almost caught the white whale. //

The units into which this passage is divided are, according to Hunt, "minimal as to length"; and "each would be grammatically capable of being terminated with a capital letter and a period" (1965, p. 21). Mean T-unit length itself is calculated by dividing the total number of words in a passage, in this case 67, by the total number of T-units, in this case 5. Thus the mean T-unit length for this passage is 13.4 words.

T-units can be described with reference to the four traditional sentence types--the simple, the compound, the complex, and the compound-complex. According to Hunt's definition of the T-unit, simple and complex sentences would each have one T-unit, but compound and compound-complex sentences would each have at least two T-units. The following sentences will illustrate.

Simple Sentences

I bought an apple. //

Having eaten my potatoes, I ate an apple. //

Complex Sentences

When I had eaten my potatoes, I ate an apple. //
I like the movie we saw about Moby Dick, the white whale. //

Compound Sentences

I ate my potatoes // and then I ate an apple. //
Jeff brought home a school assignment, // but he went to Barry's house to play. //

Compound-Complex Sentences

After I ate my potatoes, I ate an apple, // and then I went outside to play. //
The woman who had red hair broke her leg // and the black-haired man broke his arm. //

As an index of syntactic maturity or complexity, the T-unit, according to Hunt, "has the advantage of preserving all the subordination achieved by the student, and all of his coordination between words and phrases and subordinate clauses." Hunt continues:

> Slicing a theme into these units does destroy the student's coordination between main clauses, or, more accurately, between T-units. . . . But to do so is a gain rather than a loss, if one is searching for an index of maturity. For it has been shown that certain fourth graders destroy the significance of sentence length by their tendency to string T-units together endlessly with *and* after *and*, forgetting to use periods. In fact, coordination with *and*

> between T-units is an index of immaturity significant for grade at the .01 level. So such destruction is merciful. (Hunt, 1965, pp. 21-22)

Excessive coordination of independent clauses and inadequate subordination of one idea to another are, of course, two problems which also appear in the written discourse of some college freshmen, even though both problems may be less pronounced for the older group than for elementary students. Nevertheless, for a given population, the longer mean T-unit length is, the greater seems the probability that the writers have been able to subordinate less important ideas to those expressed in main or independent clauses. Another measure of syntactic maturity or complexity, the ratio of T-units to total clauses, may, however, be a better indicator of this phenomenon than mean T-unit length.

Mean Clause Length. Mean T-unit length gives some indication of how writers put together ideas in minimum sentences. Mean clause length, on the other hand, gives some indication of the relative semantic complexity or density of both dependent and independent clauses within sentences. Mean clause length is calculated by using procedures similar to those for slashing passages into T-units. If the "Moby Dick" passage, which was initially slashed into T-units, were also slashed to indicate the number of dependent or subordinate clauses, it would look like this:

> I like the movie / we saw about Moby Dick the white whale // the captain said / if you can kill the white whale Moby Dick / I will give this gold to the one / that can do it // and it is worth sixteen dollars // they tried and tried // but / while they were trying they killed a whale and used the oil for lamps // they almost caught the white whale //

To the five T-units (marked with "//") into which we initially divided the sentence, we have added five divisions to mark the dependent clauses. These divisions are marked with a "/". Thus we have divided the sentence into its ten clauses, five independent and five dependent or subordinate. By dividing the total number of words of the passage (67) by the total number of clauses (10), we are able to calculate the mean clause length for the entire passage--in this case, 6.7 words.

Subordination Index. As we have pointed out, the total number of clauses, dependent and independent, in the passage is 10 and that the total number of T-units is 5. To calculate the subordination index, the ratio of independent clauses to total clauses, we simply divide the total number of clauses by the total number of T-units. For the "Moby Dick" passage, we find that on the average each T-unit contains two clauses, one of which is an independent clause. Thus we observe a 1-to-1 ratio of independent to dependent clauses. Calculating this ratio gives us some notion of the amount of clausal subordination in a given passage or group of passages. Indirectly, and in conjunction with mean T-unit length and mean clause length, this ratio also suggests the amount or depth of embedding. That is to say, it gives some indication of the number of clauses reduced to less-than-clause status that have been embedded in both the T-units and the clauses.

Coordination-Within-Sentence Index. As we have pointed out, the lengths of sentences can vary considerably according to punctuation skill and writing style. Thus any index which is based on sentence length must be viewed with reservations. However, the amount of coordination among independent clauses does indicate something of the development of syntactic control: younger writers tend to coordinate more independent clauses within sentences than older writers do. For the "Moby Dick" passage, calculating this index allows us to quantify the obvious heavy coordination within the sentence, and the writer's lack of punctuation skill. For this passage, we find a coordination-within-sentence ratio of 5 to 1, that is, 5 T-units per sentence.

Mean Sentence Length. As should be apparent, mean sentence length tells us less precisely than either mean T-unit length, mean clause length, or mean number of clauses per T-unit anything about the syntactic abilities of writers. But sentence length does suggest some developmental trends and has been used (see Marckworth and Bell, 1967) to distinguish among many kinds of discourse produced by skilled writers, discourse ranging from mystery fiction to government documents.

It should be pointed out that these five indices relate to one another mathematically in the following way. By multiplying mean clause length by the mean number of clauses per T-unit, one arrives at the mean number of words per T-unit; and if one multiplies the mean length of T-units by the mean number of T-units per sentences, one arrives at the mean number of words per sentence. Thus the following simple formula obtains and can be used to check the accuracy of the calculations for these five indices of syntactic maturity or complexity.

$$W/C \times C/T = W/T \times T/S = W/S$$

It should, however, be pointed out that the numbers appearing in tables are frequently rounded numbers, so that the possibility of some differences does exist.

Free Modifiers. In addition to Hunt's indices of syntactic maturity or complexity is a set of measures identified by the late Francis Christensen. Christensen's work has been accessibly anthologized in Notes Toward a New Rhetoric (1979), the current edition of which lists Francis Christensen and Bonniejean Christensen as co-authors but which contains nine important essays by the former and only three non-essential, prefatory sentences by the latter. Christensen's generative rhetoric of the sentence, with its emphasis on the cumulative sentence, is important to the study of syntactic structures in written discourse. The cumulative sentence, as Christensen defines it, is a sentence whose movement from beginning to end is typified by additions to the base clause or main clause, additions which function as non-restrictive modifiers placed either before, within, or after the base clause. Christensen called these additions "free modifiers"; and in grammatical terms they can take several different forms: participial, infinitive, adjective, and prepositional phrases; absolutes and appositives; and non-restrictive clauses marked either with a subordinating conjunction or a relative pronoun.

These non-restrictive additions to base clauses can be illustrated as follows:

> Participial Phrase
>
> One of the most common free modifiers is the participial phrase, consisting of a series of words introduced by the "-ing" or "ed" form of a verb.
>
> Infinitive Phrase
>
> The principal reason for adding free modifiers to a base clause is to add information or content economically, to say more in fewer words.
>
> Adjective Phrase
>
> The free modifier, long or short, adds specificity to the base clause.
>
> Prepositional Phrase
>
> The cumulative sentence, with its free modifiers, can be quite long.

Absolute Construction

The cumulative sentence, its free modifiers often placed to the right of the verb, seems to deliver its sense as it moves along.

Appositive Construction

I like the movie we saw about Moby Dick, the white whale.

Non-Restrictive Relative Clause

Short relative clauses, which often add nonessential information to a sentence, can easily be reduced to prepositional phrases.

Non-Restrictive Clause Marked with a Subordinating Conjunction

Free modifiers, because they can significantly alter the rhythm of a sentence, are frequently used to achieve certain stylistic effects.

It is important to note that the procedures adopted for counting free modifiers demand that they be punctuated correctly, that is, set off by one or more commas from the rest of the sentence. These procedures thus make the use of free modifiers in syntactic analysis as much a test of punctuation skill as it is a test of syntactic manipulation.

Free modifiers, of the types noted, have been found to predict a significant percentage of the variance among the quality scores assigned holistically to groups of essays. Nold and Freedman (1977), for example, showed that 12% of the variance in the holistic scores assigned to essays written by Stanford Universtiy freshmen was predicted in a regression analysis by the percentage of total words in final free modifiers, those added to the ends of sentences. Faigley (1979c), using similar statistical techniques, drew similar conclusions in his study of the writing of University of North Dakota freshmen, showing that 16% of the variance among quality scores was predicted by the percentage of T-units with final free modifiers. Elsewhere, Faigley (1979a) compared the writing of his North Dakota freshmen to that anthologized in Hall and Emblen's A Writer's Reader (1976) on the basis of words appearing in initial, medial, and final free modifiers. The results of this comparison appear in Table 3.5. The fact that the figures for the Hall-Emblen writers include analyses for several kinds of discourse--narration, description, exposition, and argumentation--should indicate the importance of examining the occurrence of free modifiers in the written discourse of E 306 students.

Insert Table 3.5 about here.

When we examined the "high" and "low" quality essays of E 306 students in terms of these syntactic features--whether those identified by Christensen or those identified by Hunt--we had followed certain procedures and rules to insure consistency. These procedures and rules are those used by the Miami University investigators (Morenberg, et al., 1978) and by Faigley (1979c).

Comparison of "High" and "Low" Quality Essays. With the Hunt indices of syntactic "maturity" and the Christensen free modifiers used as the criterion variables, how do the two groups of E 306 essays compare?

First, a comparison of the two groups based on Hunt's five indices of syntactic maturity or complexity. The relevant data for this comparison are summarized in Table 3.6. As Table 3.6 indicates, the average T-unit length of the "high" quality E 306 essays is 1.64 words longer than the average T-unit of the "low" quality essays. That is to say, the writers of the "high" quality essays wrote, on the average, T-units which were a little over 12% longer than those found in the "low" quality essays. The clauses of the "high" quality essays were also larger; in fact, they were, on the average, 1.81 words or 24.2% longer than the clauses found in the "low" quality essays. As should be expected, the differences in mean T-unit length and mean clause length are reflected in mean sentence length, with the sentences of the "high" quality essays running some 6.7% longer than the sentences of the "low" quality essays. Similar differences, but in the opposite direction, also obtain for the other two indices, mean number of clauses per T-unit and mean number of T-units per sentence. The "high" quality essays exhibited, on the average, nearly 11% fewer clauses per T-unit and about 5.1% fewer T-units per sentence than did the the "low" quality essays.

Insert Table 3.6 about here.

Apart from the obvious differences in the length or number of particular features, what do these differences tell us about the writing of the two groups of students? The answer to this question must be somewhat speculative in nature because the indices used to calculate these differences actually tell us very little that will allow us to account precisely for the syntactic structures the better writers used to lengthen their T-units, clauses, and sentences and to reduce the number of clauses per T-unit and the number of T-units per sentence.

The longer T-units and, to some extent, the longer sentences may indicate that the better writers were able to control more information per independent clause than were the poorer writers.

Consider the following series of T-units, keeping in mind that the T-unit can be "immaturely" (and, in some cases, artificially) lengthened.

(1) The woman, who had red hair, fell down. (8 words)
(2) The woman having red hair fell down. (7 words)
(3) The woman with red hair fell down. (7 words)
(4) The red-headed woman fell down. (6 words)

Most would probably agree that all four of these T-units say the same thing, deliver the same content; yet they differ from one another with respect to the syntactic structure each employs to deliver a portion of that content. In question here is, of course, the embedded sentence "The woman had red hair." That embedded sentence is realized differently at the surface level in each of the T-units. In T-unit (1) the embedded sentence takes the form of a relative clause; in T-unit (2), the form of a participial phrase; in T-unit (3), the form of a prepositional phrase; and in T-unit (4), the form of a pre-nominal adjective. T-units (2), (3), and (4), all employ structures which, in effect, reduce the relative clause of T-unit (1) to less-than-clause status. Through such clause reductions, writers are able to reduce the total number of words required to deliver a given content, thus enabling them to say more in fewer words. Although Hunt's indices do not tell us whether the writers of the "low" quality E 306 essays actually used relative clauses where embedded participial or prepositional phrases or pre-nominal adjectives would have been more economical, we suspect that the better E 306 writers tended more often than did their counterparts to use the more economical embeddings and clause reductions, a suspicion partially confirmed by the fewer average number of "clauses per T-unit" observed in the "high" quality E 306 essays.

These types of clause reductions more than likely also account for the longer clauses of the "high" quality essays. The same four T-units cited above provide support for this conjecture. Consider, for example, T-unit (1). With the embedded relative clause, the mean length of the two clauses in this

T-unit is 4 words. Now consider T-units (2), (3), and (4), each of which exhibits a different clause reduction so that each T-unit contains but one clause, the independent or base clause. If we calculate the clause length for each of these three T-units, we find that it is identical to the T-unit length, or 7 words, 7 words, and 6 words, respectively. Thus our three clause reductions have the new effect of lengthening the average clause length in each case while, at the same time, reducing the length of the T-unit.

Now if we take this kind of economical clause lengthening to be typical of the writing of the better E 306 students, then we are able to account for the longer T-units of the better E 306 writers. There are primarily two ways by which a T-unit can be lengthened, either by adding subordinate clauses to it or by embedding reduced clauses into it. Since the "high" quality essays exhibit fewer clauses per T-unit than do the "low" quality essays, it is reasonable to infer that the better E 306 writers tended to embed a larger number of reduced clauses into the base or independent clause of the average T-unit than did the poorer writers. Because the clauses of the better writers tended also to be longer, it is also reasonable to infer that even in some subordinate clauses, the better writers embedded reduced-clause structures, such as prepositional and participial phrases and pre-nominal adjectives.

We have not, of course, discussed all the syntactic structures that writers have at their disposal to lengthen T-units and clauses. Nor shall we. But we will name a few of those commonly employed by college freshmen: coordinate subjects, coordinate predicates, serial constructions, ellipses, and single-word adverbs. We have, on the other hand, devoted considerable attention to several non-restrictive phrases and clauses which qualify as "free modifiers," each of which can serve to lengthen substantially a T-unit or a clause.

To the occurrence of these "free modifiers" in the "high" and "low" quality E 306 essays we now turn our attention. For each of our two groups of essays, we calculated the percentage of total words used as free modifiers and the percentage of total T-units displaying one or more free modifiers. These data are summarized as Table 3.7 and Table 3.8.

Insert Table 3.7 about here.

Insert Table 3.8 about here.

Table 3.7 and Table 3.8 reflect differences between the "high" and "low" quality E 306 essays similar to those we observed with reference to Hunt's five indices. In fact, we can assume that the differences between the two groups of essays observed with reference to Hunt's indices can be accounted for by the differences in occurrence of free modifiers in the two groups. As Table 3.7 indicates, over 10½% of the total number of words in the five "high" quality essays were employed by the better writers in "free modifiers," while only less than 6% of the total words of the "low" quality essays appeared in "free modifiers." The words in the "high" quality essays appearing in free modifiers were distributed across the three categories of "initial," "medial," and "final," with 39.1%, 15.2%, and 45.7% of the total words used as free modifiers falling into the respective categories. In contrast, of all the words appearing in free modifiers in the "low" quality essays, none was found in the "initial" position, 28.3% were found in the "medial" position, and 71.7% were found in the "final" position. Not only, then, were there fewer words, relatively speaking, in the free modifiers of the "low" quality essays than in the free modifiers of the "high" quality essays, but those words were distributed across the categories in entirely different ways. In terms of the percentage of T-units having free modifiers, the difference between the two groups of essays is equally remarkable. As Table 3.8 indicates, while 28.44% of the T-units in the "high" quality essays contained free modifiers, only 10.1% of the T-units in the "low" quality essays contained them. Differences between the two groups with reference to the distribution of free modifiers across the categories--initial, medial, and final--necessarily complement the distribution calculated previously in terms of percentages of total number of words.

What do these observed differences between the two groups of essays tell us about the writing behaviors of their authors?

Actually, they tell us a great deal. From these data on the occurrence of free modifiers, it is possible to infer that a significant portion of the difference between the two groups of essays along the dimensions of Hunt's five indices is accounted for by the greater incidence of free modifiers in the "high" quality E 306 essays. We can also infer that these differences are probably attributable to the heavier reliance, in the "high" quality E 306 essays, on relatively sophisticated syntactic structures, structures which do not appear with regularity in the discourse of less skilled writers. Table 3.5 supports this inference: it is these very structures, the ones constituting the corpus of free modifiers discussed earlier, which seem to distinguish the written discourse of accomplished, highly skilled writers from that of less accomplished, less skilled writers. As far as our two groups of E 306 writers are concerned, the writers of the "high" quality essays appear to be closer to the "norms" for the accomplished, skilled writers.

How do these University of Texas E 306 students compare with other writers, both of the same age and of different ages? As we suggested earlier, considerably more data for making such a comparison are available for Hunt's five indices than for free modifiers. In fact, it is possible to illustrate the relative standings of the E 306 students along normative or developmental continua for each of the five indices of syntactic maturity which Hunt identified. To show these standings, we have plotted the "pooled" scores of the "high" and "low" quality E 306 essays along side the scores reported for different groups of writers in other studies. These data appear in Table 3.9.

Insert Table 3.9 about here.

As Table 3.9 shows, the averages for the E 306 writers fit reasonably well into the overall pattern of development for writers ranging in age and ability from third graders to "skilled adults" who had published non-fiction professionally. For the E 306 writers, mean T-unit length and mean sentence length each exceed slightly the comparable means for Hunt's twelfth graders. But the mean clause length for the E 306 writers is somewhat less, 0.27 words, than that for Hunt's twelfth graders. The mean number of clauses per T-unit is, however, larger than that for Hunt's twelfth graders, as is the mean number of T-units per sentence. Not only do the means for the E 306 writers tend to follow developmental trends set by the younger writers studied by Hunt and O'Donnell, et al., but they also anticipate the further development along these dimensions exhibited by professional writers. The mean length of the E 306 T-units is shown to be shorter than the mean for the professional writers studied by Cooper, et al., Faigley, and Hunt. Similarly, the clauses which these same professional writers produced were longer than those produced by the E 306 writers. A less clear developmental pattern is, however, observed with respect to the mean number of clauses per T-unit, although one might argue that the level of clausal embedding within T-units reaches its maximum around the twelfth grade. The longer sentences of Hunt's professional writers not unexpectedly parallel the occurrence of longer T-units and clauses for the same group. As for the mean number of T-units per sentence, an indicator of the amount of coordination among main or independent clauses within sentences, it would seem that a good deal of such coordination appears in the written discourse of fourth graders, somewhat less in that of eighth graders, and considerably less in the writing of twelfth graders. However, beyond the twelfth grade, it would appear that writers once again begin using more coordinate main clauses, but not nearly so many as they did as eighth graders. All of this is to say that with the possible exception of mean clause length and mean number of clauses per T-unit, the writers of the "high" and "low" quality E 306 essays tend to write very much as they would be expected to write. The shorter clauses and the larger number of clauses per T-unit would, however, seem to indicate that as a group the E 306 writers are perhaps a little behind their counterparts at other universities in their ability to embed reduced clauses in the sentences they write. However, it should be pointed out that with a sample size of 10, one or two students writing significantly fewer words per clause and significantly more clauses per T-unit than the norm for his/her age group could easily skew the data.

We can also highlight a few differences between the E 306 writers and their counterparts at other universities. In comparison with Faigley's North Dakota freshmen and Stewart's

New Brunswick freshmen, the E 306 students seem to write T-units of about the length one would expect. However, the T-units of the E 306 students do seem significantly shorter than those of Cooper, et al.'s SUNY-Buffalo freshmen and Sodowsky and Witte's Oklahoma freshmen. These differences can, perhaps, be explained in terms of the apparent higher ability level of the SUNY-Buffalo freshmen and in terms of the various writing assignments used to elicit the written discourse from the various populations. As we have pointed out, a number of researchers have shown considerable variability in such length factors for discourse written in different modes; and it is known, for example, that the writing assignment used in the SUNY-Buffalo study elicited discourse of an argumentative nature while the E 306 and the North Dakota assignments encouraged students to support a central idea by drawing on their personal experience, usually expressed in the narrative and/or descriptive modes. The clauses of the pooled E 306 essays were also shorter than those found in the written discourse of the four other groups of college freshmen listed in Table 3.9. And the mean number of clauses per T-unit for the E 306 students was greater than for two other groups of college freshmen. Although we believe both findings indicate that the E 306 students should probably be reducing some of the clauses they write, we also believe that part of this difference may be attributable to the different writing assignments and part attributable to the fact that the University of Texas exempts nearly 30% of its freshman population from E 306. This exemption policy prevented us from examining the writing of the best students entering the

University of Texas as freshmen during the Fall of 1978, students whose abilities probably compare favorably to those of Stanford and SUNY-Buffalo freshmen.

In contrast to a plethora of studies using Hunt's five indices, only three studies were available that examined the occurrence of free modifiers in the writing of college freshmen; and only two of these three present their data in a form which lends them to comparison with the data for E 306 students. The relevant data from these two studies--Faigley's (1979a) study of North Dakota freshmen and Nold and Freedman's (1977) study of Stanford University freshmen--and the relevant data for the E 306 students appear in Table 3.10.

Insert Table 3.10 about here.

Table 3.10 indicates that in comparison with the University of North Dakota entering freshmen, the E 306 students use the syntactic structures which constitute the set of free modifiers at a substantially higher rate. In fact, the beginning E 306 students used 21.5% more words in final free modifiers than did their counterparts at the University of North Dakota. With regard to the percentage of total T-units which included free modifiers in the final position, the E 306 students used free modifiers in 43.3% more T-units than did the North Dakota freshmen. In comparison with the Stanford University freshmen, the E 306 students used 23.4% fewer words in medial free modifiers but 54.3% more words in final free modifiers. Because the three writing topics used to elicit the data on which these figures are based varied considerably, it is difficult to interpret the meaning of these comparative figures. However, it would seem that E 306 writers do use free modifiers at a rate which appears not to be below the rate of use by other beginning college freshmen. Nevertheless, one look at Table 3.5 would suggest that E 306 writers have a long way to go before the frequency of occurrence of free modifiers in their writing approaches that of the professional writers anthologized in Hall and Emblen's A Writer's Reader (1976).

Although the foregoing analysis of syntax in the written discourse of 10 E 306 students is by no means an exhaustive one nor a statistically sophisticated one, we hope that it does present data which will enable sound decisions with respect to the kinds of instruction these students seem to need in their E 306 classes. We, however, will not discuss pedagogical strategies until we have presented all of our analyses of the "high" and "low" quality E 306 papers and the results of the five principal and the five derivative comparisons.

3.3 The Occurrence of First-Person and Second-Person Pronouns in "High" and "Low" Quality Essays

In the previous sections we distinguished between the "high and "low" E 306 essays on the basis of error frequencies and syntactic complexity. In this brief section, we will present another distinction, this one based on the frequency of occurrence of first-person and second-person pronouns. On the basis of Lunsford's analysis of essays written for the British Columbia English Placement test (1979), we anticipated that E 306 essays with low holistic scores would use more first- and second-person pronouns than essays with high holistic scores.

The relevant data for this comparison are summarized in Table 3.11 and Table 3.12. As Table 3.11 indicates, first-person pronouns constitute about one-half of one percent of the total words in the "high" quality E 306 essays and nearly one percent of the total words in the "low" quality essays. Converted into frequency of occurrence, these percentages mean, according to Table 3.12, that the writers of the "high" quality essays used some form of the first-person pronoun once every 201.75 words or once every 13.19 T-units, while the writers of the "low" quality essays used a first-person pronoun once every 112.67 words or once every 8.25 T-units. Thus in terms of word-frequency only, the writers of the "low" quality E 306 essays used a first-person pronoun approximately 79% more often than did the writers of the "high" quality essays. The differences between

---

Insert Table 3.11 about here.

Insert Table 3.12 about here.

---

these two groups of E 306 essays are even more striking in terms of the occurrence of second-person pronouns. Table 3.11 indicates that of all the words in the "high" quality essays, only 0.06 of one percent of them were second person pronouns, while 3.55% of the words in the "low" quality essays fell into that category. Table 3.12 expresses this noteworthy difference between the two groups in two additional ways. For the "high" quality essays, a second-person pronoun appears once every 1,614 words or once every 105.5 T-units; but for the "low" quality essays, a second-person pronoun appears once every 28.17 words or once every 2.06 T-units. In terms of the frequency of occurrence in words, this means that the writers of the "low" quality essays used some form of the second-person pronoun 49 times as often as did the writers of the "high" quality essays.

These differences are also reflected in the combined occurrence of first-person pronouns and second-person pronouns in the two groups of essays. According to Table 3.11, less than six-tenths of one percent of the words in the "high" quality essays were either first- or second-person pronouns, while nearly 4½% of the words in the "low" quality essays were. This difference is similarly reflected in Table 3.12. In the "high" quality essays, a first- or second-person pronoun appeared once every 197.33 words or once every 11.72 T-units. In contrast, in the "low" quality essays, some form of the first- or second-person pronoun appeared once every 22.53 words or once every 1.65 T-units. Thus in terms of the frequency of occurrence in words, the writers of the "low" quality essays used a first- or second-person pronoun almost eight times as often as did the writers of the "high" quality essays.

We do not know, of course, whether the occurrence of first- and second-person pronouns is a predictor of writing quality, in these ten essays or in essays of other college freshmen. That could only be established through the analysis of a sample considerably larger than the 10 essays we analyzed and through the application of sophisticated regression analyses of the resultant raw data. However, these differences between the "high" and "low" quality essays may be compatible with what other researchers and theorists have had to say about the writer-audience relationship in written discourse, a relationship which has been a major concern to rhetoricians from classical antiquity to the present. Modern-day discourse theorists--whether interested principally in poetics as in the case of Jakobson (1960) or in rhetoric as in the case of Moffett (1968) or Britton (1975)--generally agree with the concepts of discourse component interaction expressed in Kinneavy's (1971) communication triangle. Kinneavy holds that the function of a piece of written discourse is determined by the primary emphasis of the piece, on the writer (self-expression), the audience (persuasion), the subject matter (reference discourse), or the medium (literary). Moffett treats the functions of discourse in terms of the possible interactions which can exist among the "I" (writer), the "you" (audience), and the "it" (the subject matter). Moffett's discussion these interactions in terms of the three classes of personal pronouns in the English language is germane to this analysis.

Britton adds another dimension, both implicit and explicit in Kinneavy's model, that of physical and psychological distance between the writer and his/her audience. Without developing our notion to the extent it deserves, we hypothesize that the greater occurrence of first- and second-person pronouns in the "low" quality E 306 essays on this topic suggests that these writers, and we are here drawing on Britton's Piaget-informed analyses of the writing of British schoolchildren, have not yet developed the ability to abstract themselves from either their audiences or their subject matters. That is to say, the occurrence of first- and second-person pronouns in the "low" quality E 306 essays suggests that the world view of these students is still very much an egocentric one. This is not to imply anything denegrating about these writers; it is only to suggest that psychologically they have yet to develop as much as their counterparts who wrote the "high" quality essays. Given the time and the proper encouragement, their writing will become less egocentric when occassion demands, just as surely as their T-units and clauses will grow longer as they, the students, grow older.

3.4 Text Cohesion

Each of the three previous analyses has provided insight into textual differences between student essays rated high and low in quality, but none indicates that we will be able to characterize adequately the quality of essays on the basis of features at and below the sentence level. While the sentence may be the basic unit of communication, it is important in written texts only insofar as it relates to the context of the whole discourse.

If repetition is any measure of the importance a particular group attaches to a given concept, then the problems associated with coherence in the writing of college freshmen represent one of the major concerns of composition teachers and textbook writers alike. To some extent, this concern with coherence results from inadequate materials and methods for teaching it; and to some extent, it results from a general ignorance of the many devices writers have at their disposal for achieving coherence in extended discourse.

It seems to us that the state of the art is partly attributable to the considerable attention composition researchers and teachers devote to matters which can be discussed almost exclusively in terms of sentence-level operations and features. Punctuation and grammar instruction, which often warrant a good deal of time in the classroom, is focused within the boundaries of a sentence, unless the teacher not intelligently regards paragraphing as a punctuation skill. Similarly, matters of style are most often treated with reference to sentence-level syntactic structures or lexical choices within sentences --all this in spite of the composition teacher's goal of teaching students to write units of discourse extending beyond the sentence, and even beyond the paragraph.

When coherence is taught, it is usually taught in one of three ways: by presenting examples for the purpose of distinguishing between the concept of unity and the concept of coherence; by imploring students to use traditional transitional words such as conjunctive adverbs between their sentences; and by impressing upon student writers the importance of thesis statements, topic sentences, logical patterns, development of ideas, support, specific details, etc. Of course, all of these strategies and concepts are to one degree or another useful, because they approach the idea of text from the standpoint of a semantic unit, as an extended series of meanings whose components must somehow be made to relate to one another. Yet problems of coherence persist in the writing of many college freshmen, primarily because (we think) such strategies and concepts as those detailed above, while at some point generally useful and usefully applicable, do not point students and their teachers to specific devices they need to achieve or to teach coherence in an extended discourse.

Others have, of course, recognized the inadequacies of these traditional concepts and strategies. And several recent composition researchers have formulated alternative concepts and strategies. A few examples, not an exhaustive survey, follow. During the mid-1960's, Christensen (1965) extended his notion of a generative rhetoric of the sentence to include a generative rhetoric of the paragraph. Christensen maintained that when beginning transitional sentences are excluded, most paragraphs move from a general semantic level to increasingly more specific semantic levels. Thus following the most general sentence--the "topic" sentence--subsequent sentences exist at a lower level of generality relative to the "topic" sentence. Within a given paragraph, a non-topic sentence could thus be coordinate or subordinate in level of generality to the preceding sentence. Here is an example of such a paragraph, its topic sentence labeled "(1)" to indicate the higher level of generality and the subsequent sentences marked with larger numbers to show their lower levels of generality relative to the topic sentence.

(1)
(2)
(2)
(3)
(4)
(3)
(3)

All sentences not labeled "(1)" exist at a lower level of semantic generality than sentence (1); that is to say, they are subordinate in meaning to sentence (1). In this sequence the sentences labeled "(2)" are coordinate in meaning to one another, whereas the sentences labeled "(3)" and "(4)" are subordinate in meaning to the sentences labeled "(2)." And so on. Grady (1971), D'Angelo (1974), and Wittig (1975) have extended to the essay Christensen's idea of semantic levels of generality in the sentence and the paragraph.

While Grady, D'Angelo, and Wittig make important applications of Christensen's theory to larger units of discourse, these writers do not actually advance the theory itself. However, Karrfalt (1968) and Pitkin (1969) do. Karrfalt, taking issue with Christensen's contention that paragraphs progress to successively lower levels of generality, observes that some sentences in some paragraphs extend the meaning of a previous sentence to a higher level of generality. In this, Karrfalt anticipates Pitkin's concept of superordination. Pitkin identifies, on the basis of function within the whole discourse, hierarchically ordered semantic "blocs" of thought, which may well extend beyond the boundaries of paragraphs but which do not necessarily extend beyond the boundaries of, say, a subordinate clause in a given sentence. Pitkin argues that within the context of the whole discourse, these "blocs" are related to one another through coordination, through complementation (as in a cause-to-effect pattern), through subordination (as in a classification pattern where the first "bloc" presents the distinguishing features of a class and a second "bloc" specifies members of that class), or through superordination (the opposite of subordination and where the semantic movement is, for example, from the specific instances to the general class). Pitkin's theory of discourse "blocs" is compatible with that of Rodgers (1966, 1967) who notes, quite rightly, that "paragraphs are not composed; they are discovered" and that a given semantic unit of discourse may span several semantic chunks indented as paragraphs.

Also important to discovering principles of coherence and of arrangement in texts are the ideas advanced by the tagmemicists. Drawing on the linguistic theories of Kenneth Pike (1964a, 1964b, 1965, 1967), Becker and Young have studied paragraph structure in terms of certain functional semantic slots which they see as constituting semantic patterns when they are combined in paragraphs. Becker (1965) proposes three such patterns, each consisting of two or more functional slots --the QA (question-answer), the PS (problem-solution), and the TRI (topic-restriction-illustration)--and the various possible combinations and permutations of these three basic types. The underlying theory and the implications of this system for analyzing paragraphs are elaborated by Young and Becker (1965) elsewhere.

Now spanning several disciplines, research in text comprehension and production has established that syntactic structures are formed relatively late in production and forgotten early in comprehension. Consequently, much empirical research in discourse comprehension and many theoretical models of text structure follow semantic, rather than syntactic, lines; and several of the models now being tested empirically view texts as underlying sets of semantic propositions arranged hierarchically (e.g., Frederiksen, 1977; Kintsch, 1974; Meyer, 1975). Considerable empirical evidence indicates that sentences located high in the structural hierarchy of a text are recalled more frequently than sentences located low in the heirarchy (Kintsch 1974; Kintsch & Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Meyer, 1975; Thorndyke, 1977).

As noted above, researchers in written composition have recognized the importance of semantic relationships in extended units of discourse, and some have suggested ways such relationships might be analyzed (Winterowd, 1970; Cummings, Herum, & Lybert, 1971). These systems, however, are difficult to apply empirically to individual texts. More recently, European text linguists began considering the question of what is a text. One criticism of early text structure models advanced by linguists (e.g., Harris, 1952) was that a satisfactory definition of a text was assumed by such models but did not exist in fact.

Halliday and Hasan address the problem of textuality in Cohesion in English (1976). They set out a systematic and presumably comprehensive method for analyzing cohesive ties in extended discourse. They do not attempt to provide a model of text structure but to define a text in terms of cohesive relationships among sentences. Their system for analyzing cohesive ties in a stretch of discourse considers a text as a semantic unit, linked together by cohesive ties. Cohesion, therefore, "is a semantic relation between an element in a text and some other element that is crucial to the interpretation of it" (1976, p. 8). Central to Halliday and Hasan are the notions of presupposing and presupposed elements; and although presupposing and presupposed elements can co-occur within a sentence, it is cohesive ties between presupposing and presupposed elements across "sentence boundaries" which allow sequences of sentences to be realized as a text unit.

Halliday and Hasan's system for analyzing and classifying cohesive devices provides for five major classes of cohesive devices, nineteen subclasses, and numerous sub-subclasses.

Here we will be concerned with the five major classes of cohesive devices--Reference, Substitution, Ellipsis, Conjunction, and Lexical Reiteration and Collocation--and their respective subclasses.

Cohesion through Reference. In English there are a large number of words of various kinds which in a text "make reference to something else for their interpretation" (p. 31) rather than "being interpreted semantically in their own right" (p. 31). These words are "directives indicating that information is to be retrieved from elsewhere" (p. 31). The information which is retrieved "is the referential meaning, the identity of the particular thing or class of things that is being referred to; and the cohesion lies in the continuity of reference, whereby the same thing enters into the discourse a second time" (p. 31). Reference is, however, not exclusively concerned with items in the text itself: the presupposed item can exist outside the text proper in the situational context. For example, it is possible to utter the following sentence in a given situational context and expect the listener to understand fully what is meant: "Will you turn that off?" Both you and that presuppose two objects which do not exist in the text itself. Similarly, in written discourse we frequently encounter the first-person plural pronoun, as in this sentence, or the singular form or second-person pronouns; and we do so without having the presupposed specific person or persons named in the text itself. In the previous sentence, we could have one of any number of possible antecedents outside the text itself, but it has none which is text explicit. Reference beyond the text itself to the context of the discourse is "exophoric" reference; and according to Halliday and Hasan, exophoric reference "makes no contribution to the cohesion of a text" (p. 53). Again according to Halliday and Hasan, certain occurrences of _you_, _one_, _we_, _they_, and _it_ simply make it possible for a writer "to conform to the structural requirements of the clause, which demands a nominal in various places" (p. 53). Reference within the text itself is called "endophoric" and may be either "anaphoric," referring to some previous item or items in the text, or "cataphoric," referring to some item or items which follow in the text. These two types of endophoric reference may be illustrated as follows:

Anaphoric Reference

(1) John bought an apple. //*

(2) He bought it yesterday. //

Cataphoric Reference

(3) That was good. //

(4) It was the best meal I've ever eaten. //

*Here and in all subsequent examples the double slashes represent T-unit boundaries. This practice is employed because text cohesion deals, technically, with discourse across T-units and because the T-unit will be one of the denominators used later in this section to quantify cohesion in the E 306 essays.

Sentence (2) contains two examples of anaphoric reference: _He_ and _it_ depend for their meaning on the prior occurrence of _John_ and _apple_ in sentence (1). Sentence (3) and sentence (4) each contain an example of cataphoric reference, although in treating matters of text cohesion we are not concerned with within-sentence reference as in sentence (4). Nevertheless, the demonstrative that in sentence (3) looks forward, rather than backward, to meal for its meaning. Anaphoric and cataphoric cohesive ties can arise for all three of the subclasses of reference cohesion--cohesion through pronominals, cohesion through demonstratives and definite articles, and cohesion through comparatives.

The most common type of reference cohesion is that which is achieved through the use of personal pronouns. The example of anaphoric reference in sentences (1) and (2) above contains two cohesive ties based on the appearance of personal pronouns: the pronouns he and it in sentence (2) take their meaning from the nouns John and apple in sentence (1). Another example, drawn from the set of E 306 essays we examined is the following:

Reference Cohesion (Pronominal)

(5) At home, my father is himself. //

(6) He relaxes and acts in his normal manner. //

Across-sentence or across-T-unit cohesion here, as in sentences (1) and (2), results from the pronouns He and his of sentence (6) presupposing and gathering up into themselves the meaning of father in sentence (5). Sentence (6), without sentence (5), would be impossible to process, for without the presence of the presupposed item, father, no cohesive tie would exist.

A second type of reference cohesion is that achieved through the use of demonstrative pronouns or definite articles. Our example of cataphoric cohesion supplies one example of the use of a demonstrative pronoun to realize a cohesive tie, with That of sentence (3) looking forward to the occurrence of meal in sentence (4). One of the most common uses in the E 306 essays of demonstratives to achieve cohesion is illustrated by sentences (7) and (8).

Reference Cohesion (Demonstratives)

(7) We question why they tell us to do things. //

(8) This is part of growing up. . . . //

The appearance of This in sentences (8) presupposes information retrievable from sentence (7), without which sentence (8) would have no meaning of its own. It is perhaps too obvious a point to make, but notice that we are not able to process the meaning of we, they, and us in sentence (7) because the presupposed items are no where in the text provided, that is, no cohesive ties exist. The use of the definite article can also create a cohesive tie within a text similar to the ones created through the use of demonstratives. Consider, for example, the following pair of sentences taken from an E 306 paper.

Reference Cohesion (Definite Article)

(9) Humans have many needs, both physical and intangible. //

(10) It is easy to see the physical needs such as food and shelter. //

The in sentence (10) functions very much as the demonstrative those would have functioned had it appeared in the place of the. The definite article in sentence (10) provides a cohesive tie by virtue of the fact that the noun it modifies, in this case needs, has a specific referent in the preceding sentence (9). The definite article in sentence (10) helps specify the noun needs in the same sentence, thus making needs in sentence (10) more specific than needs in sentence (9); and in this way the anticipates the appearance of food and shelter at the end of sentence (10). Only in its presupposing function, however, does the constitute a cohesive tie.

A third type of reference cohesion is achieved through the use of comparatives. Comparatives become cohesive ties within texts by virtue of the impossibility of writing a sentence such as "X is like" and expecting meaning to result: we must always write "X is like Y." As with pronouns, when the "something else" exists outside the text, the reference is an exophoric one and not cohesive; when in the text, it is endophoric and cohesive. Comparatives need not, however, name a particular "Y": they may stress "likeness and unlikeness, without respect to any particular property" (p.77). Such comparisons are said to be "general." In contrast, "particular comparison" involves comparison in terms of specific quantities or qualities. General comparison typically involves comparison focusing on identity, similarity, or difference between two "things." Particular comparison, on the other hand, usually specifies "likeness" or "unlikeness" in terms of quantitative modifiers such as better or equally good. It would, of course, be tedious to illustrate all possible types of comparative reference which establish cohesion, but one example from an E 306 essay will suggest the importance of cohesive ties based on comparatives.

Reference Cohesion (Comparatives)

(11) The males ruled the women; //

(12) the women obeyed and formed themselves into the mold the males constructed for them. //

(13) But today it is different. //

In this T-unit sequence, the word different, a general comparative, presupposes and gathers up into itself the sense of the past expressed in sentences (11) and (12). Without sentences (11) and (12), it would be impossible to understand sentence (13). Although different is not a personal or demonstrative pronoun or a definite article, it does refer anaphorically to the meaning expressed in sentences (11) and (12), meaning which is presupposed by it. Thus different in sentence (13) constitutes a cohesive tie, intimating a contrast between the past and the present.

Cohesion Through Substitution. Whereas reference cohesion is achieved in texts by meaning-based relationships across sentence boundaries, cohesion through substitution is achieved through relationships based on wording. Mush less common in texts than reference cohesion, cohesion through substitution refers simply to the "replacement of one item by another" (p. 88). Cohesion through substitution can take three forms, involving, respectively, replacement of *nominals*, *verbals*, or *clauses* by other items.

Cohesion through nominal substitution is illustrated in sentences (14) and (15).

Cohesion Through Substitution (Nominal)

(14) I have yet another behavior for the classroom. //

(15) That *one* society or, in this case, the university deems appropriate for the classroom. //

The word *one* in sentence (15) replaces or substitutes for the nominal *behavior* in sentence (14). The nominal *behavior* could very easily have appeared in sentence (15) without having created an unduly awkward expression; and herein lies an important difference between reference cohesion and cohesion through substitution. Often the presupposed item in a cohesive tie based on reference cannot be substituted for the item doing the presupposing without sacrificing, for lack of a better word, the "style" of the sequences. Consider, for example, sentences (5) and (6) above. The *he* and the *his* in sentence (6) cannot be replaced by the presupposed *father* and [*father's*] in sentence (5) without making the entire sentence sequence terribly awkward. Substitution, on the other hand, seems merely a matter of economy, and the cohesive tie which results from substitution (especially for nominals) can easily be achieved through cohesion based on lexical repetition, which is discussed below.

A second type of cohesion through substitution involves verbs or verbals. Consider sentences (16) and (17).

Cohesion Through Substitution (Verbal)

(16) A father may expect his son to develop interest and skill in playing football, //

(17) but the son may not do so. //

In this pair of T-units, do so in sentence (17) substitutes for or replaces to develop interest and skill in playing football in sentence (16). In this latter T-unit all of the meaning of the earlier infinitive phrase is gathered up into its replacement or substitute, do so. The appearance of do so in T-unit (17) is not a matter of ellipsis, which is also a type of cohesion, because nothing in the previous infinitive phrase can be added grammatically to do so, whereas ellipsis is always "substitution by zero".

The final type of cohesion through substitution focuses on the replacement of clausal structures by a word or words not found in the clause being replaced. Again, the example is drawn from an E 306 essay.

Cohesion Through Substitution (Clausal)

(18) Do all people change their behaviors this easily? //

(19) It seems so.

In the case of this pair of T-units, so replaces an entire clause without renaming any element in the clause itself. Sentence (19) literally means "It seems [that all people do change their behavior this easily]," a meaning which can be understood from sentence (19) only because an anaphoric cohesive tie exists between the clausal substitute so in sentence (19) and the only clause in sentence (18).

Cohesion Through Ellipsis. Cohesion through ellipsis is, as we have already intimated, in some respects similar to cohesion through substitution. And, as we have also noted, it may even be called "substitution by zero" (p. 143). Yet the cohesive ties that result from each differ. In cohesion through substitution, a given word or set of words presuppose and replace the presupposed element. In ellipsis cohesion, the presupposing structure gives the impression of having left something unsaid while, at the same time, often repeating some element from that which is presupposed. For example, in the sequence "Who is taller? John" the cohesive tie is realized because is taller is ellided in the second part of the sequence, the interrogative who serving as a cataphoric reference for John. In cohesion through substitution, as in sentence (19), a particular word, so, fills a particular structural slot in the sentence and serves as a "place-marker for what is presupposed, whereas in ellipsis nothing is inserted in the slot" (p. 143). As with substitution, there are three subclasses of cohesion through ellipsis--nominal ellipsis, verbal ellipsis, and clausal ellipsis.

Cohesion through nominal ellipsis can be understood with reference to a presupposed nominal, the presupposing of which is frequently a modifying element for the nominal. In nominal ellipsis the nominal itself is ellided and the modifying element changes its function within the presupposing sentence. A good example of exophoric nominal ellipsis appeared repeatedly in the speech of post-WWII West Germans, who referred to Konrad Adenauer as "Der Alte," a term of affection which ellides "Mann" of "Der alte Mann." An excellent example of endophoric nominal ellipsis appeared in one of the E 306 essays we examined:

Cohesion through Ellipsis (Nominal)

(20) Among most teenagers we can observe behavioral patterns that fall into three categories. //

(21) These might be called the good, the bad, and the ugly. //

In sentence (21) good, bad, and ugly are all instances of a nominal having been ellided by the addition, in the presupposing sentence, of modifying elements; and in the resulting ellipsis of behavioral patterns there exist three cohesive ties with the previous sentence. Thus in nominal ellipsis further specification of the ellided nominal can easily be achieved. A case could also be made for the plural demonstrative these in sentence (21) as supplying a cohesive tie through ellipsis of the nominal categories in sentence (20). Demonstratives, as we illustrated with sentences (7) and (8), often create cohesive ties through anaphoric reference. However, for the demonstrative this in sentence (8), it is not possible to supply the presupposed item in the form of a single-word nominal or by substituting one or ones for it, whereas such is possible for these in sentence (21) as in these categories. Hence this in sentence (8) represents a cohesive tie achieved through reference, while these in sentence (21) represents a cohesive tie achieved through ellipsis.

Cohesion can also be achieved through verbal ellipsis. Verbal ellipsis also involves "substitution by zero." In verbal ellipsis, a presupposed verb phrase or portion of a verb phrase does not fill the expected slot in the presupposing sentence. Sentences (22) and (23), from an E 306 essay, provide an example of a cohesive tie based on verbal ellipsis.

Cohesion through Ellipsis (Verbal)

(22) Do these changes in roles point to insincerity within us? //

(23) To a degree, they do. //

In this example of verbal ellipsis, sentence (23) supplies a part, do, of the verb phrase do . . . point to insincerity within us, the remainder of the verb phrase being understood from the preceding sentence. Hence the ellided part of the verb phrase supplies a cohesive tie between sentences (22) and (23).

A third type of ellipsis is clausal ellipsis where an entire clause from a preceding sentence is ellided from a presupposing sentence. As with nominal and verbal ellipsis, the resulting cohesive tie is anaphoric and involves "substitution by zero." Sentences (24) and (25), drawn from an E 306 essay, will illustrate:

Cohesion through Ellipsis (Clausal)

(24) Can role-playing be reduced? //

(25) Yes, if our society becomes more liberal. //

In this sequence, yes in sentence (25) provides a cohesive tie by elliding the clause role-playing can be reduced which is expressed as an interrogative in sentence (24). Note, however, that with the presupposing yes and similar such words, no part of the presupposed clause is repeated.

Cohesion Through Conjunction. The fourth major class of cohesive devices is that known as conjunction. Conjunction differs substantially from cohesion through reference, substitution, and ellipsis--all of which depend heavily on anaphoric relations to create cohesive ties. The principal difference between conjunction and the other types of cohesion we have discussed are summarized by Halliday and Hasan:

> Conjunctive elements are cohesive not in themselves but indirectly, by virtue of their specific meanings; they are not primarily devices for reaching out into the preceding (or following) text, but they express certain meanings which presuppose the presence of other components in the discourse. (p. 226)

There are five major types of conjunction which establish cohesive ties across sentences, between presupposing and presupposed items in the text. These are additive conjunction, adversative conjunction, causal conjunction, temporal conjunction, and continuative conjunction. Because the cohesive function for each of these types is basically the same, we will provide successive examples of all types. Then we will discuss the examples collectively, rather than individually. All of these examples come from E 306 essays we examined.

Cohesive Conjunction (Additive)

(26) The college student who wore faded blue jeans now wears three-piece suits, //

(27) and he feels the need to be approved of and accepted by his new associates. //

Cohesive Conjunction (Adversative)

(28) People can usually be themselves when they are with their friends. //

(29) However, people who are attempting to make friends often feel a need to impress these prospective friends. //

Cohesive Conjunction (Causal)

(30) Today's society sets the standards. //

(31) The people more or less follow it [sic]. //

(32) Consequently, there exists the right behavior for the specific situation at hand. //

Cohesive Conjunction (Temporal)

(33) A friend of mine went to an out-of-state college, //

(34) and before she left, she expressed her feelings about playing roles to win new friends. //

Cohesive Conjunction (Continuative)

(35) Noone [sic] wants to be rejected, //

(36) and to prevent rejection we, of course, change our behavior often. //

As should be clear from our set of examples, both coordinating conjunctions--such as *and*, *but*, *nor*, or *yet*, and *so*--and conjunctive adverbs--such as *however*, *consequently*, and *moreover*--can serve conjunctive cohesion by supplying cohesive ties across sentence boundaries. In connection with the types of words, especially coordinating conjunctions, that can supply cohesive ties, it should be noted that conjunction, as far as cohesion is concerned, is not the same as coordination. With coordination there is the built-in notion of something like equivalence of semantic stature, as in the case of coordinate subjects or coordinate predicates. With cohesive conjunction the underlying notion is simply that of joining together explicitly for some particular purpose.

It is also important to note that a cohesive relationship can exist between two sentneces without a conjunction marker present. For example, consider the following sentence sequence from John Koffend's "The Case for Alcohol" (*Atlantic Monthly*, Dec., 1979, p. 68).

(37) Drink was so routine an embellishment of American colonial life that the Puritans called it "the good creature of God" and arrived with plenty of spirits aboard the *Mayflower*. //

(38) Cups brimming with wine fermented from local grapes were emptied at the first Thanksgiving. //

(39) So insistent was the forefatherly thirst that when local crops of wine grapes, malt, and hops proved inadequate, our resourceful ancesters

produced potent skullpoppers from artichokes, cornstalks, pumpkins, turnips, and horseradish. //

(40) Casks were broached as a matter of course at ministerial investitures; //

(41) many of the clerics were themselves champion boozers, and did not shrink from admitting it. //

(42) Wage contracts of the period frequently ratified . . . the human right to inhale an occasional skinful.

In this sentence sequence, there appears no conjunctive marker between sentences (40), (41), and (42). Yet no one would probably dispute that between these T-units and exists in "spirit," although not in "letter." Sentence (41) quite clearly is additive, as is (42), both adding examples to the one expressed in T-unit (40), examples which illustrate the importance of alcohol during the colonial period. The coherence or cohesion which exists among the T-units, rather than depending on specific conjunctions, is a direct result of the semantic relationship of these T-units with sentence (37). This brings us to an important point, namely, that when analyzing conjunctive cohesive ties, we are concerned only with those which are marked explicitly in the text.

Turning to our examples, then, we observe that between T-units (26) and (27) there exists a text-explicit conjunction between the two T-units, even though one might question whether the additive and joining the two T-units is to be preferred over the subordinating conjunction because. The examples of adversative, causal, and temporal conjunction are fairly straightforward. Each of the cohesive markers presupposes the meaning contained in a previous sentence or set of sentences.

A word about continuative conjunction, however, may be useful. The class of continuatives contains such items as now, of course, well, anyway, surely, and after all used in a cohesive way. These words function as cohesive devices when they carry the force of their non-reduced forms. Of course, which appears in sentence (36) actually means you should have known that already, that ought to be obvious to you, or this is to be expected. According to Halliday and Hasan, such a continuative becomes cohesive "when it is slipped in as an incidental or an an afterthought, since its interpretation becomes contingent on the context (and therefore on the preceding text)" (p. 271.).

Lexical Cohesion. The fifth and last major kind of text cohesion is that which includes cohesive ties based on lexical relations. Lexical cohesion is realized through the selection of vocabulary; and it is somewhat more difficult to identify than the other four types of cohesion because every lexical item has the potential for entering into a cohesive relationship and because there is never anything inherent in the occurrence of a given lexical item that makes it cohesive. If we were to encounter the word this in a text, we would automatically refer anaphorically to the text in order to recover the meaning of the demonstrative. Similarly, if we were to encounter one as a nominal substitute or an adjective filling the functional slot of a nominal, we would, for the one, look back into the text to determine what had been replaced by one and, for the other, look back at the text to determine which nominal had been ellided. So, too, with conjunction. When we encounter a conjunctive adverb such as however, we attempt to establish from the preceding text the nature of the adversative relationship introduced. Lexical cohesion, in contrast, results from the non-random occurrence of lexical items in the text, from the "patterned occurrence of lexical items," from the movement of the text "from one topic to another. . . on reasonably systematic lines with a certain consistency of topic and predictability of development" (p. 288).

Perhaps the most important notion associated with lexical cohesion is that of "environment." Each lexical item in a text has an environment "that will provide the context within which the item will be incarnated on . . . [a] particular occasion" (p. 289); and within this within-text environment or context of each lexical item, the "instantial" or "text meaning" of the item is determined, "a meaning which is unique to each specific instance" (p. 289). Our processing of a text involves our becoming aware of the "environment" of any lexical item appearing in the text so that "by the time any given lexical item is taken in, its context has already been prepared. The preceding lexical environment is perhaps the most significant component of this context," providing "a great deal of hidden information that is relevant to the interpretation of the item concerned" (p. 289). Although the lexical environment of a given lexical item includes not only words which are relevant to it semantically but also all words in the preceding text, "it is the occurrence of the item in the context of related items that provides cohesion and gives to the passage the quality of text" (p. 289). In determining lexical cohesion, "relatedness is a matter of more or less; there is no clearly defined cutoff point such that we can say that sunset, for example, is related to just . . . [one] set of words and no others; and it is the closeness of the relationship that determines the cohesive effect" (p. 289).

There are "two kinds of relatedness, one kind being relatedness in the linguistic system and the other being relatedness in the text" (p. 290). Halliday and Hasan elaborate:

> There are degrees of proximity in the lexical system, a function of the relative probability with which one word tends to co-occur with another. Secondly, in the text there is a relatedness of another kind, relative proximity in the simple sense of distance separating one item from another, the number of words or clauses or sentences in between. The cohesive force that is exerted between any pair of lexical items in a passage of discourse is a function of their relative proximity in these two respects. (p. 290)

Cohesion is greatly influenced by the "overall frequency" of a given lexical item "in the system of the language" (p. 290). As a general rule, the higher the "overall frequency" of a given lexical item in the language, "the smaller the part it plays in lexical cohesion in texts" (p. 290). Thus in analyzing a text for lexical cohesion, one must employ a great deal of common sense, relying on internalized knowledge "of the nature and structure [of the language's] vocabulary" (p. 290) and on one's judgment as to what "constitutes a significant pattern and what does not" (p. 290). Obviously, in determining lexical cohesive ties within a text, we would ignore repetition of grammatical items--such as pronouns, high frequency verbs, verbal auxilliaries, and prepositions--since these can occur with the full range of lexical items in the language.

Lexical cohesion can be achieved in a text through reiteration and through collocation. In what follows we treat lexical collocation and four types of lexical reiteration, reiteration by means of the same item, of a synonym or near-synonym, of a superordinate, and of a general item.

Between the types of lexical reiteration, there is no clear-cut distinction, the result being a "continuum" of lexical reiteration ranging from repetition of the same item, to repetition by the use of a synonym or near synonym, to repetition through the use of a superordinate, to repetition through the use of a general item, and, even, to repetition through pronominal reference (see the section Reference Cohesion above). This continuum can be illustrated by examining two sentences which occurred sequentially in an E 306 essay and by examining some possible alternative renderings of the same two sentences.

(43) Drinking is typical of many college freshmen, particularly members of some fraternities. //

(44) This behavior is usually accompanied by obscene talk. //

Although this is not a particularly striking set of sentences, it does illustrate how a superordinate such as behavior in sentence (44) depends for its "instantial" meaning on the prior appearance of a related and more specific lexical item, in this case the gerund *drinking* in sentence (43). In short, *behavior* occasions an anaphoric cohesive tie between sentence (44) and sentence (43). This pair of sentences, together with some possible revisions, will also illustrate the other types of lexical reiteration and their relationship to reference cohesion.

Drinking is typical of college freshmen. . . . //

The drinking
The boozing
The behavior] is usually accompanied by obscene talk. //
The thing
It

With this illustration, we can see how a lexical cohesive tie between two sentences can be created through the repetition of the same item, in this case *drinking*. Similarly, the use of *boozing*, a synonym, establishes a cohesive tie across the boundaries of the two sentences. In both cases, part of the meaning of the second item, whether it is the same item or a synonym, must be recovered from the earlier occurrence. A somewhat different type of lexical cohesion between the two sentences is created if *behavior* is used as the presupposing item since *behavior*, in this case a superordinate, has the capacity to refer to any number of different activities, only one of which is drinking. Having even less specificity is the general noun, *thing*, which can include within its domain things such as automobiles and sweaters, as well as activities such as drinking. Lastly, the personal pronoun *it* can also be used, illustrating the point at which lexical and reference cohesion merge. It should also be noted that the in four of the five instances cited has a cohesive function, discussed above in connection with reference cohesion. Finally, we should point out that in these same four cases, the demonstrative this can be used in place of the, thus creating an alternative set of cohesive ties.

Each of these four types of lexical reiteration are illustrated in the following sentence sequences derived from E 306 essays we examined. The items creating the cohesive tie in each pair are underlined.

Lexical Reiteration (Same Item)

(45) Human behavior varies greatly, depending on the situations one experiences. //

(46) One's behavior can change due to the people he surrounds himself with. //

Lexical Reiteration (Synonym or Near-Synonym)

(47) One behaves according to the expectations of those around him. //

(48) Obviously, a young adult acts differently around friends than around his parents. //

Lexical Reiteration (Superordinate)

(49) Some professional tennis players, for example, grandstand, using obscene gestures and language to call attention to themselves. //

(50) Other Professional athletes do similar things, such as spiking a football in the end zone, to attract attention. //

Lexical Reiteration (General Item)

(51) Some professional tennis players, for example, grandstand, using obscene gestures and language to call attention to themselves. //

(52) Other professional athletes do similar things, such as spiking a football in the end zone, to attract attention. //

All the lexical cohesive relationships which cannot be properly subsumed under lexical reiteration are included in a miscellaneous class called collocation. Collocation refers to lexical cohesion "that is achieved through the association of lexical items that regularly co-occur" (p. 284). Lexical cohesion through collocation is perhaps the most difficult type of cohesion to analyze in a given text because items said to collocate involve neither repetition, synonymy, superordination, nor general items. Consider, for example, the following two T-units:

(53) The eggs were delicious. //

(54) The bacon was crisp. //

Lexical cohesion obtains across these two T-units by virtue of two words, eggs and bacon; but these involve no repetition and no synonymy; neither is one a superordinate for the other; in context neither is one a general item which can include within its semantic domain both the class of items represented by the other and classes of items not included in the class of the other.

Yet the two words are systematically related to one another, primarily because the two words are frequently used together in the language. That is to say, they collocate. In Halliday and Hasan's words, a collocative cohesive tie exists

> between any pair of lexical items that stand to each other in some recognizable lexiosemantic (word meaning) relation. This would include not only synonyms and near-synonyms such as *climb* . . . *ascent*, *beam* . . . *rafter*, *disease* . . . *illness*, and superordinates such as *elm* . . . *tree*, *boy* . . . *child*, *skip* . . . *play*, but also pairs of opposites of various kinds, complementaries such as *boy* . . . *girl*, *stand up* . . . *sit down*, antonyms such as *like* . . . *hate*, *wet* . . . *dry*, *crowded* . . . *deserted*, and converses such as *order* . . . *obey*.
>
> It also includes pairs of words drawn from the same ordered series. For example, if *Tuesday* occurs in one sentence and *Thursday* in another, the effect will be cohesive; similarly *dollar* . . . *cent*, *north* . . . *south*, *colonel* . . . *brigadier*. Likewise with any pairs drawn from unordered lexical sets, like *basement* . . . *roof*, *road* . . . *rail*, *red* . . . *green*. The members of such sets often stand in some recognizable semantic relation to one another; they may be related as part to whole, like *car* . . . *brake*, *box* . . . *lid*, or as part to part, like *mouth* . . . *chin*, *verse* . . . *chorus* (or *refrain*); they may be co-hyponyms of the same superordinate term, *i.e.*, both members of the same

> more general class, such as *chair* . . . *table* (both hyponyms of *furniture*), *walk* . . . *drive* (both hyponyms of *go*); and so on. (p.285)

The members of such sets as these somehow relate to one another semantically; and while it may be possible, although certainly not easy, to describe in some detail the nature of these relations, it is not necessary to do so for the purpose of identifying cohesive collocational ties in a text or set of texts. What is important is that the items said to collocate "share the same lexical environment" (p. 286).

Three T-unit sequences from the E 306 papers we examined will help illustrate the cohesive effects of *collocation* in written discourse, one set of collocating items being underscored in each sequence.

(55) *Parents* expect certain kinds of behavior. //
(56) *Children*, however, cannot always live up to these expectations. //

(57) Around *adults*, we usually put on our good behavior masks. //
(58) Instead of replying "yeh," a *teenager* would reply "Yes, sir." //

(59) On a *camping trip* with their parents, teenagers behave more responsibly than they do at home. //

> (60) Unwilling to help with household chores, at the camp site they gather wood for fire, carry tents and sleeping bags over long distances, and even get water from a creek or lake. //

As should be apparent from these examples, it is possible to have rather long chains of collocational cohesion, chains which can span across several T-units rather than only two as in the examples. The examples also point out that collocational ties "do not depend on referential identity and are not of the form of reiteration accompanied by the or a demonstrative--in other words, all lexical cohesion that is not covered by . . . 'reiteration.'"

Comparison of "High" and "Low" Quality Essays.

Using a coding system derived from the one developed by Halliday and Hasan (see pp. 333-355), we analyzed our sets of five of the best and five of the poorest E 306 essays (as determined by the holistic evaluations of pretest essays written on the "Changes in Behavior" topic) in terms of the five major types of cohesion--reference, substitution, ellipsis, conjunction, lexical--and the various subclasses specified above for each major type. We hypothesized that the qualitative differences between the two sets of five essays each would be reflected in the number and kinds of cohesive ties employed by the two groups of writers respectively.

In order to compare the two essay sets along these 24 dimensions of text cohesion, we calculated for each dimension for each essay group five indices: (1) the mean frequency in words of cohesive ties, determined by dividing the total number of words in a given essay set by the total number of cohesive ties of each type and subtype (W / CT Type) in that set; (2) the mean number of cohesive ties per T-unit, calculated by dividing the total number of T-units in a given essay set into the total number of cohesive ties of each type and subtype in that set (CT Type / T); (3) the percentage of the total number of cohesive ties represented by each type and subtype of cohesive ties in an essay set (CT Type / CT); (4) the number of cohesive ties per 100 T-units for a given essay set, which index is (CT Type/ T) X 100; and (5) the ratio of cohesive ties to total words in an essay set, a ratio which can be expressed as a percentage by multiplying the ratio by 100.

A comparison of the "high" and "low" quality E 306 essays follows, beginning with reference cohesion, proceeding through lexical cohesion to a comparison on the basis of total cohesive devices, and ending with a brief comparison of E 306 writers with SUNY-Buffalo writers follows.

As was demonstrated above, reference cohesion takes three forms--cohesion by means personal pronouns, demonstratives and the cohesive the, and comparatives. The relevant data for this comparison are summarized in Table 3.13.

Insert Table 3.13 about here.

The four indices listed across the top of Table 3.13 give us four ways of comparing the occurrences of referential cohesive ties in the two sets of E 306 essays. As Table 3.13 indicates, some form of reference cohesion occurs once every 24.96 words in the "low" quality essays and once every 19.4 words in the "high" quality essays. The difference between the two word frequencies indicates that one must read 5.56 more words in the "low" quality essays before encountering a referential cohesive tie than in the "high" quality essays. Table 3.13 also reveals that a substantial portion of this difference can be attributed to the higher frequency of personal pronouns in the better essays. Personal pronouns as cohesive ties appear with a frequency nearly four times as great in the best essays as in the worst, with one personal pronoun appearing once every 63.4 words in the best essays and once every 233 words in the worst. Although the difference in frequency of occurrence of personal pronouns is the most pronounced, the occurrence of demonstratives and cohesive definite articles is also important, with one of these appearing approximately every 78 words in the poorer essays and one every 59 words in the better essays. A marked difference also shows up between the frequencies of occurrence of comparatives, this time the higher frequency belonging to the "low" quality essays. While the higher frequency of personal pronouns, demonstratives, and definite articles is to be expected in the better essays, which have more content to tie together than their counterparts, the lower frequency of comparatives was not. We had expected that the presumably more sophisticated writers, those receiving "high" holistic scores, would show a greater tendency to compare concepts, objects, actions, and persons represented in the texts. For these groups of essays on this particular writing topic, such was not, however, the case.

The differences between the two essay sets along the dimensions of reference cohesion are also reflected, perhaps more markedly, by other indices included in Table 3.13. Because the number of cohesive ties per 100 T-units is simply the index CT Type / T-units multiplied by 100, we will focus on the two indices listed at the far right of Table 3.13, the percentage of total cohesive ties and the mean number of cohesive ties per 100 T-units. The latter first. As the figures in Table 3.13 indicate, the "high" quality essays exhibited nearly twice as many referential cohesive ties per 100 T-units as did the "low" quality essays, 84 per 100 T-units as opposed to 48, or 0.841 referential cohesive ties per T-unit compared to 0.478 per T-unit. This difference is, again, largely attributable to the greater use of personal pronouns by the better writers, who used 5 times as many personal pronouns per 100 T-units as did the poorer writers. Demonstratives and definite articles as cohesive ties appear almost twice as often per 100 T-units in the better essays than they do in the poorer essays, while the poorer essays display about 3.6 more comparatives per 100 T-units than the better essays. Whether one uses the word as the unit of measurement or the T-unit, differences between the two groups clearly obtain. Analyses based on the T-unit may, however, be misleading, unless one keeps in mind that the mean word length of the T-units in the better papers is 1.64 words longer than that for the poorer papers (see Table 3.6 above).

Cohesive ties resulting from nominal, verbal, and clausal ellipsis and substitution account for the fewest cohesive ties in both the best and the worst E 306 essays. For the best essays, 2.49% of all cohesive ties were the result of the three types of substitution and 2.68% the result of the three types of ellipsis. In contrast, 1.4% of all cohesive ties in the "low" quality essays were the result of substitution, all of these being nominal substitutions; and 4.2% of all cohesive ties resulted from ellipsis, principally nominal ellipsis but including a few instances of verbal ellipsis. The relevant data for substitution and ellipsis cohesion are summarized in Table 3.14 and Table 3.15, respectively.

Insert Table 3.14 about here.

Insert Table 3.15 about here.

Substitution and ellipsis represent relatively sophisticated ways of achieving cohesion across T-unit boundaries. Thus one would expect that the better E 306 writers would have employed significantly more such cohesive ties than did the poorer writers. As Table 3.14 indicates, some form of substitution occurred once every 126.9 words in the best essays and once every 349.5 words in the "low" quality essays. In terms of cohesive ties per 100 T-units, this means that 12.9 substitution ties per 100 T-units appeared in the "high" quality essays but only 3.4 in the "low" quality essays, or that substitution occurred about 3.6 times more often in the better essays than in the poorer essays. It should be reiterated that the better writers used all types of substitution while their counterparts used only nominal substitution. For cohesion achieved through ellipsis, the differences between the two groups are, in general, less pronounced, the poorer writers producing 10.2 such ties per 100 T-units as compared with 13.8 such ties for the better writers. However, while the better writers used all three types of ellipsis, the poorer writers employed only two, nominal and verbal. Yet the relative frequency in words of ellipsis ties is essentially the same for both groups, an elliptical tie occurring once every 116.5 words in the poorer essays and once every 117.8 words in the better essays. In general, our survey of cohesive ties based on ellipsis and substitution within the two essay sets suggests that the better writers tended to employ substitution by zero and substitution by not-zero more readily than did their counterparts.

Just as differences between the two essay sets obtain for reference, substitution, and ellipsis cohesion, so too do differences based on the occurrence of the five types of conjunctive ties--additive, adversative, causal, temporal, and continuative. The relevant data for our two sets of E 306 essays are summarized in Table 3.16.

Insert Table 3.16 about here.

As Table 3.16 indicates, the frequency in words of all conjunctive ties in the better essays is more than twice as large as that for the poorer essays. For the better essays, one conjunctive tie occurs once every 25 words, but for the poorer essays a conjunctive tie appears only about once every 58 words. This substantial difference in the frequency of conjunctive ties is attributable to the fact that four of the five types of conjunctive cohesion appear with a significantly greater frequency in the better essays than in the poorer essays. Indeed, two of the types which figure importantly in the better essays appear not at all in the poorer essays. These differences are most pronounced when they are viewed in terms of occurrences per 100 T-units. Additive conjunctive ties occurred more than twice as often per 100 T-units in the better essays than in the poorer essays, 22.8 compared with 10.2. Adversative ties occurred 9.9 times per 100 T-units in the better essays but not at all in the poorer essays. Whereas causal conjunctions appeared 13.9 times per 100 T-units in the "high" quality essays, they occurred only 1.7 times per 100 T-units in the "low" quality essays. Temporal ties appeared about the same number of times per 100 T-units in the two essay sets, 8.5 in the "low" and 7.9 in the "high." As with adversative conjunctions, no continuatives appeared in the poorer essays, but were used nearly 11 times per 100 T-units in the better essays. Collectively, conjunctive ties appeared in the "low" quality essays 20.4 times per 100 T-units, but they appeared 65.4 times per 100 T-units in the "high" quality essays, or more than three times as often. These differences are similarly reflected in the percentage of total cohesive ties accounted for by the five types of conjunctive cohesion. In the better essays, conjunctive ties accounted for 12.64% of all cohesive ties in that set, but they accounted for only 8.4% in the set of poorer essays.

Thus far we have summarized the differences between the two sets of E 306 essays along the dimensions of cohesion realized through grammatical relations--references, substitution, and ellipsis--and through conjunction, which bridges the gap between grammatical and lexical cohesion. Now we turn our attention to a comparison of the two essays sets based exclusively on lexical cohesion, that is, on reiteration and collocation. It will be remembered that lexical reiteration consists of four subtypes--reiteration through repetition of the same item, reiteration through synonymy, reiteration through superordination, and reiteration through the use of general items. As Table 3.17 illustrates, lexical cohesion accounts for most of the cohesive ties

Insert Table 3.17 about here.

found in both essay sets, accounting for 66.43% of all ties in the poorer essays and 65.9% in the better essays. The slight difference between percentages suggests a greater similarity between the two essay sets than there, in fact, is. In terms of the frequency in words of the occurrences of lexical ties, we observe that the poorer writers created a lexical tie every 7.36 words while the better writers managed such a tie every 4.79 words. This rather large difference, 2.57 words, between the mean distances separating lexical ties in the two essay sets is explained by the frequencies for the types of lexical rieteration and for lexical collocation. The mean distance separating same items in the better essays is on the average 2.16 words shorter than that in the poorer essays; and the mean distance separating synonyms is 24.6 words shorter. Superordination was not used at all by the poorer writers but was used once every 274.83 words by the better writers. General items also appear more often in the better essays than in the poorer, once every 126.85 words as compared to once every 174.75 words. So too with the occurrence of collocational ties: these appear once every 41.12 words in the poorer essays and once every 17.36 words in the better essays. These differences are even more noticeable when they are viewed in terms of their occurrence per 100 T-units. Over a text span of 100 T-units, a same-item tie occurs 105.1 times in the poorer essays and 179.2 in the better essays. Synonymy appears 20.3 times and 48.5 times per 100 T-units for the two groups, respectively; and nearly twice as many general items appear in the better essays as in the poorer essays per 100 T-units. One of the most striking differences is the occurrence of collocational ties, with 94.1 per 100 T-units appearing in the better essays but only 28.8 in the poorer essays.

The percentage of total lexical ties represented by the four types of lexical repetition and by collocation is also revealing. These percentages appear in Table 3.18.

Insert Table 3.18 about here.

To some extent, these percentages suggest a greater range of cohesion-achieving skill among the writers of the "high" quality essays. Especially important in this regard seem to be the differences between the two groups in the use of same-item repetition, synonymy, and collocation.

These differences between the two sets of E 306 essays can also be viewed collectively. When, for example, we divide the total number of words in a given essay set by the total number of cohesive ties observed in that set, we find that on the average one encounters in the poorer essays a cohesive tie of some type once every 4.881 words and in the better essays a cohesive tie once every 3.159 words, a difference in excess of 1.7 words. A comparable difference obtains for the mean number of cohesive ties per T-unit in the two essay sets. On the average, 2.4237 cohesive ties appeared in each T-unit of the poorer essays, while 5.1693 ties per T-unit appeared in the better essays. This is to say that the better E 306 writers established, on the average, over twice as many cohesive ties per T-unit as did their counterparts. In short, the better writers produced discourse which, from the standpoint of cohesion, was considerably more dense than that of their poorer counterparts. This cohesive density is similarly reflected in the total number of cohesive ties per 100 T-units for each essay set. In the poorer essays, 242.4 cohesive ties occurred for each 100 T-units; but over twice that many, 516.8, appeared per 100 T-units in the "high" quality essays. Another kind of difference between the two essay sets appears when the ratio of cohesive ties to total number of words is calculated. These ratios, together with their conversions to percentages, appear in Table 3.19 for all cohesive ties and for each major type of cohesion for each essay set.

Insert Table 3.19 about here.

Table 3.19 actually requires no interpretation since it is simply another means of viewing the same data for the two essay sets. It should, however, be noted that it would be incorrect to infer from Table 3.19 that, for example, fully 31.66% of the words in the better essays function as cohesive ties. Such an inference would be valid only if all cohesive ties consisted of but one word. Some cohesive ties involve more than one word, as the following T-unit sequence illustrates.

> (61) Human behavior varies greatly according
> to the situation one is experiencing
> at the time. //
> . . . . . . [19 intervening T-units] . . . . .

(62) Thus human behavior depends on the people with whom one associates. //

In this sequence, the text cohesion realized through the repetition of the phrase human behavior would be counted as but one cohesive tie, even though four words are involved.

This T-unit sequence suggests yet another way of comparing the cohesive features of the two sets of E 306 essays, one which focuses on the distance separating the presupposing and the presupposed items in a cohesive tie and one which Halliday and Hasan describe (pp. 329-333) and employ analytically (pp. 339-355). As T-units (61) and (62) illustrate, the presupposed item in a cohesive tie does not necessarily reside in the preceding T-unit. A cohesive tie can, in fact, span several T-units before the presupposing and the presupposed items actually enter into a cohesive relationship. And the span of text crossed by a given cohesive tie provides the basis for this other way of classifying text cohesion. It would be possible, for example, for the pronominal it to appear in the first of the 19 T-units, which we shall label (61A), deleted from the above sequence. In that case human behavior of T-unit (62) would be mediated through the referential it back to the presupposed human behavior in T-unit (61). If a second hypothetical T-unit, which we label (61B), follows (61A) and includes the phrase human behavior, then a different kind of relationship obtains. The following modification of the earlier T-unit sequence will help illustrate the points we are trying to make.

(61) Human behavior . . . . //

(61A) It . . . . //

(61B) . . . human behavior . . . . //

. . . . [N intervening T-units . . . .

(61P) Childish behavior . . . . //

. . . . [N intervening T-units . . . .

(62) . . . human behavior . . . . //

Notice that It in T-unit (61A) depends for its meaning on human behavior in the immediately preceding T-unit, (61). The cohesive relationship between the presupposing It and the presupposed human behavior can thus be characterized as an immediate tie. Now consider the cohesion which exists among T-units (61), (61A), and (61B). The phrase human behavior in (61B) is a repetition of the same phrase in T-unit (61), but that repetition is mediated through the It in (61A). Thus the cohesive relation between T-units (61) and (61B), which is realized through the repetition of the same lexical item, is said to be mediated by virtue of the intervening referential It in T-unit (61A). A third type of cohesive tie is that represented by the relationship of T-units (61) and (62), before we added T-units (61A), (61B), and (61P). When T-units (61) and (62) are separated by intervening T-units exhibiting no mediating cohesive ties, the cohesive relationship between the two is said to be remote, that is to say, spanning one or more T-units in which no related cohesive ties appear. A fourth type of cohesion results from the combination of mediated and remote ties. This type is illustrated by T-units (61B), (61P), and (62). In this sequence, the presupposed item, *human behavior* of (61B), is both remote from the presupposing item of T-unit (62) *and* mediated through *childish behavior* of (61P). Thus cohesive ties can be classified as *immediate*, *remote*, *mediated*, or *mediated and remote*. Unlike the earlier classifications of cohesive ties--which are based on grammatical, lexical-grammatical, and lexical cohesion--the classification now proposed is based rather on the *span of text* separating the items in a cohesive tie.

For our two sets of E 306 essays, we classified all cohesive ties according to this "span of text" system. Then we calculated the percentage of total cohesive ties represented by the four different "span of text" categories for each of the two essay sets. The results of these computations appear in *Table 3.20*.

Insert *Table 3.20* about here.

As *Table 3.20* indicates, both groups of E 306 writers used about the same relative percentage of remote cohesive ties, 29.69% for the poorer writers and 25.94% for the better writers. However, the better writers used a substantially higher percentage, 41.59% to 32.80%, of immediate ties and a substantially smaller percentage, 25.86% to 36.72%, of remote-mediated ties. Mediated ties appeared in the better essays with about eight times the frequency that they appeared in the poorer essays.

The percentages listed in *Table 3.20* allow us to begin focusing, in summary fashion, on the crucial differences between the two groups of E 306 writers, differences indicated by our analyses of text cohesion. The higher relative percentage of immediate cohesive ties in the better essays suggests, among other things, that the better writers tended to establish stronger cohesive bonds between individual T-units than did the poorer writers. This conclusion finds a good deal of support in our previous analyses of reference and conjunctive cohesion in the two essay sets. It will be recalled from Table 3.1. that the better writers employed reference cohesion about twice as often, 84.1 times to 47.8 times, in a text span of 100 T-units as did the poorer writers. Because the most significant difference in the occurrence of referential cohesion between the two essay sets was located in the significantly higher frequency in the better essays of personal pronouns (25.1 per 100 T-units as compared to 5.1 for the poorer writers) and because personal pronouns do not usually span intervening T-units, we can infer that the better writers paid considerably more attention to elaborating in subsequent and immediate T-units items introduced in a given T-unit. Also figuring importantly in the greater use of immediate cohesive ties is the frequency with which the better writers employed T-unit-bridging conjunctions. It is probably accurate to say that conjunctive ties regularly create immediate cohesive ties between T-units in succession. It is not, then, surprising to find, according to Table 3.16 that the better writers used over three times as many conjunctive ties, 65.40 as compared to 20.4, per text span of 100 T-units as did the poorer writers. Not insignificant in this regard is the fact that, as Table 3.16

shows, the better writers employed all types of conjunctive cohesion while the poorer writers managed to use only three, and one of those only rarely. Thus it would appear that the better E 306 writers tended to extend the lexical-semantic domain established in a given T-unit to an immediately following T-unit considerably more often than did the poorer writers. One major effect of such lexical-semantic extensions is, of course, essay length; and it is not surprising that the essays of the better writers are, on the average, 375.2 words longer than those of their poorer counterparts. Although the typically American notion that "more" is "better" can be challenged on any number of fronts, it would appear that the longer essays enjoyed a greater probability of being rated "high" than did the shorter essays, at least as long as the "more" was linked through the use of explicit cohesive ties.

As we pointed out above, the relative percentage of mediated ties is roughly eight times as large for the better essays than for the poorer ones. This finding further supports the notion of extending the lexical-semantic domain of a given T-unit. Meidated ties, in contrast to immediate ties, enable a writer to extend the lexical-semantic domain of a given T-unit into not merely a second succeeding T-unit but into several succeeding T-units. Judging from the relative percentage with which mediated ties occurred in the two essay sets, we must conclude that the better E 306 writers can create such ties with greater facility than the poorer writers can.

The greater relative percentage of remote-mediated cohesive ties for the poorer writers may, at first, seem to contradict the trend which seems to be developing. And to some extent it does. But the higher relative percentage, 36.72% as compared to 25.86%, of remote-mediated ties for the poorer writers suggests to us an important and fundamental difference between the two group of writers. This higher relative percentage may very well complement that group's failure to elaborate and extend lexical-semantic domains through successive T-units and may well suggest that less "really" new information or semantic content was introduced during the course of a poorer essay than during the course of a better essay. That is to say, the poorer writers may have tended more toward reiteration of previously cued information than did the better writers. Indeed, in reading the poorer essays one could not help but note what might be called conceptual and lexical redundancy. Although from the standpoint of text cohesion a certain amount of redundancy is a virtue, the "redundancy" in the poorer E 306 may be a flaw, because the poorer writers provided relatively infrequent occasions for immediate and mediated cohesive ties.

The relative percentages of remote ties tends to reinforce this interpretation. Although the relative percentages do not differ substantially, 29.69 % for the poorer writers and 26.94 % for the better writers, they do suggest that the poorer writers tended more often to reach back across several T-units in order to recover lexical items. This, coupled with the smaller relative percentage of immediate and mediated ties and the higher relative percentage of remote-mediated ties, may again suggest a greater reliance on a superficial redundancy of concept and lexical item in the "low" quality essays than in the "high quality essays.

Since cohesion analysis of written discourse, student-produced or otherwise, has not the history and tradition of syntactic analyses, it is difficult to say what precisely the implications of the present analysis are for the teaching of writing. In addition, the size of our sample is far too small to allow us to generalize from it with confidence. Yet our analysis does point to certain implications. (1) Without entering into disputes focused on the chronology of the chicken and the egg, we suggest that sophistication in the use of cohesive devices may be directly tied to students' abilities to employ heuristic devices as they explore a given writing topic. That is to say, students with highly developed invention skills have a greater chance of writing better discourse from the standpoint of cohesion than students who do not have such skills. (2) Insofar as cohesion is a theoretical construct having several components addressing different kinds of semantic relationships in texts, it is ipso facto a theory of invention itself and can thus be used as a discovery heuristic. (3) Because cohesion addresses semantic relationships within texts, its major components can and should be used to guide certain of the composing processes of student writers. (4) Because the focus of cohesion is on the types of semantic relationships that obtain across T-units, it can and should be used as a basis for studying the composing and comprehension processes of language users.

Relating the present study of cohesion in written texts of University of Texas freshmen writers to patterns of cohesion in the writing of other freshmen is no easy matter, because only one such study has been done and in that relatively little information is provided. But in that study (Cooper, et al., 1980) appear a number of figures which allow us to draw some comparisons between the groups of E 306 writers we examined and the group of SUNY-Buffalo students which Cooper, et. al.., examined. These comparisons, however, are based on writing samples of two different kinds, and we suspect that the frequencies and types of cohesion differ according to discourse kind.

Some differences between the two groups of freshmen--the UT freshmen and the SUNY-Buffalo freshmen--with respect to the use of cohesive ties can be seen in general terms, while others can be seen in specific terms.

At the most general level of analysis, the cohesive density of the essays of the two freshman groups does not appear remarkably different. The average number of cohesive ties per 100 T-units for the SUNY-Buffalo population was 488 for the "high" quality papers and 280 for the "low" quality papers. For the UT freshmen about 517 and about 242 cohesive ties per same text span were observed for the "high" and "low" quality groups respectively. While these differences may in part be attributable to variation in mean T-unit length and do not appear to us particularly remarkable, we should emphasize that the "low" quality UT papers exhibited fewer cohesive ties per 100 T-units than did the "low" SUNY-Buffalo papers, while the reverse was true for the "high" quality papers.

At a more specific level, it is possible to say--on the basis of the available SUNY-Buffalo data--that the "high" quality UT papers used nearly eight times as many comparative ties as the "high" quality SUNY-Buffalo papers. This large difference is, however, probably an artifact of the different writing topics--narrative and descriptive on the one hand and argumentative on the other. Niether set of "high" quality papers relied heavily on substitution or ellipsis to achieve cohesion, although the "high" quality UT papers used both with about six times the frequency as the "high" quality SUNY-Buffalo papers. This difference may also be attributed to writing topic variation. Also likely attributable to topic variation are the large differences between the two sets of essays along the dimension of total conjunctive ties. The writers of the "low" quality SUNY-Buffalo papers used 16 such ties per 100 T-units, while the writers of the "high" quality papers used about 17. However, for the UT essay set, about 20 and 65 conjunctive ties were observed in the "low" and "high" papers respectively.

The "high" and "low" quality papers of the two populations also differed with respect to their reliance on lexical cohesive ties. While same-item lexical ties constitute about 65% and 50% of all lexical ties for the"low" and "high" quality papers, respectively, for both freshman populations, differences between the two populations are most obvious along the dimension of lexical collocation. Whereas the percent of total lexical ties represented by collocation in the "low" quality SUNY-Buffalo papers was 32%, in the "low" quality UT papers only about 20% of the total lexical ties were collocative. Similarly, whereas collocative ties represented 41% of all lexical ties in the SUNY-Buffalo papers rated "high," only about 32% of all lexical ties in the "high" quality UT papers were collocative. While these differences may simply be attributable to the respective writing topics, we suggest that the greater relative use of collocation in the SUNY-Buffalo papers may reflect the generally higher verbal abilities of the SUNY-Buffalo students from whom the writing samples were taken.

3.5 Errors, Syntax, Personal Pronouns, and Cohesion in E 306 Essays: A Conclusion

The analyses of certain features of "high" and "low" quality E 306 essays point to some major differences in the writing behaviors of the two groups of students. These analyses also suggest that in many ways the writing skills of E 306 students are as well developed or better developed than the skills of college freshmen at other institutions. Certainly the gross measurements of syntax suggest that the UT students write with a command of syntax comparable to that of students at other institutions. The more specific analyses of syntactic features, however, probably indicates that at least for the writers of the "low" quality papers instruction in the use of final nonrestrictive modifiers would be productive. With respect to error frequencies, the UT students in general appear superior to certain other groups of college freshmen for whom data were available, suggesting that greater attention in E 306 classes to error reduction may not be necessary. This suggestion finds support in the analyses of language mechanics presented in Chapter 4 and Chapter 5 of this report. The differences between the uses of first- and second-person pronouns in the "high" and "low" quality essays suggests that E 306 students, especially the writers of the "low" quality papers, need help developing a sharper sense of the relationships that exist among writer, audience, and subject matter in written discourse. Finally, our analyses--of essay length and cohesion--suggests that the poorer E 306 writers need systematic instruction in invention, a skill which, as Chapter 4 and Chapter 5 suggest, may be best taught in the "synthetic" curriculum.

CHAPTER 4

E 306 CHANGE SCORES AND FIVE DERIVATIVE COMPARISONS

In the first chapter of this report, we identified the general purpose of the present study and, in so doing, indicated the major differences between the two principal E 306 options we wished to compare. In Chapter 2, we introduced the research design, a design which would allow a large number of curricular and instructional variables to be controlled and thereby studied systematically. In Chapter 2, we also outlined the procedures employed for collecting and analyzing data in the form of essays and in the form of responses to objective tests; and we specified the sample population studied. In the third chapter, we examined in considerable detail some of the differences between "high quality" and "low quality" E 306 essays written on the A_1 writing topic (see Chapter 2). In this, the fourth chapter, we analyze and compare a number of curricular and instructional components of the two major E 306 options.

Our purpose in the present chapter is thus a dual one: (1) to determine, as far as possible with data at hand, whether E 306 is an educationally effective course in writing and (2) to begin to identify those curricular and instructional components which appear most efficacious in teaching composition to

E 306 students. To achieve this dual purpose, we will, first, examine the changes across time observed in the writing and in the writing-related behaviors of all 180 students we selected from the 20 sections of E 306 used in the present study. Second, we will evaluate different E 306 components by comparing the instructional and curricular options previously specified in Table 2.2. These analyses will provide part of the basis for a set of recommendations for improving writing instruction in E 306 (see Chapter 8).

4.1 Changes in Writing and Writing-Related Behaviors

Although the relative effectiveness of different instructional and curricular variables raises questions both interesting and important, a necessary prior question focuses on the combined effectiveness of all instructional and curricular options for E 306 which contribute to the principal and derivative comparisons specified in Table 2.1 and Table 2.2. The present section describes the pooled data for all 20 E 306 sections studied, focusing on the changes which occurred across one semester of composition instruction. First, we will treat the changes observed between the two pretest essays and the two posttest essays collected from each of the 180 students selected for study. Second, we will examine changes in writing-related skills and abilities, as observed across the same period and as measured by responses to objective tests.

Changes Identified by the Pre- and Posttest Essays. As we pointed out in Chapter 2, the 180 students chosen for the present study each completed two pretest essays and two posttest essays. One pretest essay and one posttest essay were written on the A_1B_1 topics presented in Chapter 2. Half of the students wrote on the A_1 topic and half on the B_1 topic as a pretest; for the complementary posttest essays, the topics were simply reversed. One pretest and one posttest essay were also written on the A_2B_2 topics, which were also presented in Chapter 2. Adminstration of these topics followed the same procedures used in the administration of the A_1B_1 topics.

As indicated in Chapter 2 and Chapter 3, these four essays from each of the 180 students were analyzed for total number of words, mean T-unit length, mean clause length, the ratio of T-units with final free modifiers to the total number of T-units, and the ratio of words in final free modifiers to the total number of words. When the decimal is moved two places to the left, these two ratios become percentages. In Chapter 3 a history of the four syntactic indices was presented, together with rationales for their use in the present study. Also as stated in Chapter 2, the essays were submitted to single-impression holistic evaluations by experienced teachers of college writing. The descriptive data which derive from the word-length analyses, the syntactic analyses, and the holistic evaluations of the 720 essays and the results of paired T-tests of the changes across time are reported in Table 4.1.

Insert Table 4.1 about here.

As Table 4.1 indicates, the means for the total number of words and for the syntactic indices on the A_1B_1 pretest and the A_1B_1 posttest essays do not differ significantly, the computed t-values failing to reach statistical significance at the .05 level. However, the means for the corresponding holistic rating, which represent the means for the two summed ratings (see Chapter 2), differ significantly beyond the .001 level of probability. Thus even though none of the length features of the essays elicited by the A_1B_1 topics changed significantly over a semester, the overall quality of the writing did. This latter change finds corroboration in Table 2.3, which lists the distribution of holistic scores across the four scoring categories, and Table 2.4, which lists the percentage of scores falling in each scoring category. As Table 2.3 and Table 2.4 illustrate, over 10% fewer scores of "1" and "2" were awarded the posttest A_1B_1 essays than were awarded the pretest A_1B_1 essays, while approximately 3% more scores of "3" and 8% more scores of "4" were awarded the posttest A_1B_1 essays than the pretest essays on the same topics. These higher frequencies of posttest scores falling into the two higher scoring categories account for the 0.49 improvement in writing quality over the course of the semester.

While significant change between the pretest-posttest means for A_1B_1 essays was observed only for the holistic ratings given the essays, on the A_2B_2 topics differences between the pre- and posttest means along three of the six dimensions are statistically significant--average word length of the essays, the ratio of total words used in final free modifiers to the total number or words, and mean holistic rating. The ratio of T-units containing final free modifiers to total T-units approached statistical significance but did not reach the .05 level of confidence. As with the difference between the pre- and posttest means for T-unit length and clause length for the A_1B_1 essays, the difference between the pre- and posttest means was too small to reach statistical significance.

The descriptive data and the statistical tests presented in Table 4.1, even though they do not address the question of the relative effectiveness of one or another E 306 option, are important for a number of reasons.

First, and certainly most important, these data and statistical tests strongly suggest that taken as a unit the 20 E 306 sections studied represent 20 educationally valid classes: the overall quality of the writing of E 306 students is demonstrated to have improved significantly over the period of one semester of college writing instruction. Given two holistic ratings ranging from 1 to 4 which were summed to produce a possible scoring range from 2 to 8 for each essay, Table 4.1 indicates that on the average each A_1B_1 posttest essay was one-half of a point better in quality than its pretest counterpart. The average observed change from the A_2B_2 pretest essays to the corresponding posttest essays was even larger, with each A_2B_2 posttest essay nearly three-quarters of a point better than its pretest counterpart. Because the principal goal of E 306 is to teach students to write better and because there is no reason to believe that the 180 students we studied were any better as writers at the beginning of the semester than E 306 students in general, these highly significant gains in essay quality suggest that E 306 classes in general achieve their major goal. And it is important to note in this connection that although positive change in essay quality would have been acceptable at the .05 level of confidence, the positive change in quality for both the A_1B_1 and the A_2B_2 is significant beyond the .001 level.

The descriptive data reported in Table 4.1 are important for yet another reason: they strongly suggest that the two sets of essay topics actually did elicit essays which differ in kind from each other. This suggestion finds support when the mean T-unit lengths and the mean clause lengths of the essays based on the A_1B_1 topics are compared to the corresponding means of the essays based on the A_2B_2 topics. Such a comparison reveals that the A_1B_1 topics elicited T-units on the pretest essays that were on the average 1.70 words shorter than those elicited on the pretest essay written for the A_2B_2 topics. A somewhat smaller difference--1.56 words--obtains between the mean T-unit lengths of the posttest essays. A similar difference is also observed between the mean clause lengths elicited by the two sets of essay topics: the mean clause lengths of the pretest essays differ by 1.62 words, and the means for the posttest essays differ by 1.37 words. As was pointed out earlier, the A_1B_1 topics were those used in the Miami University sentence-combining experiment and were "devised" specifically in "invite descriptive and narrative supporting details" (Morenberg, et al., 1978, p. 246). The A_2B_2 topics, on the other hand, were designed by Cameron and Witte to elicit argumentative discourse from the students who were involved in the present study. That differences along the dimensions of mean T-unit length and mean clause length should obtain across the two sets of writing topics strongly suggests that the two sets of topics did in fact elicit written discourse of two different kinds. This suggestion draws on the findings of several previously published studies. Witte and Davis (1980a) have recently pointed out that such syntactic differences among discourses written in different discourse modes "have been known to exist at least since Frogner (1933) and Seegars (1933), whose findings have been, to some extent, confirmed . . . by Kincaid (1953), Marckworth and Bell (1967), Johnson (1967), Bortz (1969), Veal and Tillman (1971), Perron (1977), and Crowhurst (1978)." Many of the more recent studies (e.g., Perron's and Crowhurst's) have used mean T-unit length and mean clause length to distinguish discourse written in different modes, typically the four traditional modes--narration, description, exposition, and argumentation. Research in this area points to the conclusion that argumentative discourse employs longer clauses and T-units than discourse written in the other three traditional modes. Certainly, the clauses and the T-units elicited by the A_2B_2 topics in the present study are appreciably longer than those elicited by the A_1B_1 topics, thus suggesting that the kinds of discourse the topics were supposed to elicit from the students were in fact elicited.

The data reported in Table 4.1 for mean clause length and mean T-unit length are also important because they allow us to compare the syntactic development of beginning E 306 students to that of other college students and to that of writers of different ages. A similar comparison, based on the five essays rated "high" in quality and the five essays rated "low" in quality was made in Chapter 3; here, however, with an N of 180 instead of an N of 10 and with reference to both sets of essays instead of one type of essay from the A_1B_1 essays, we offer a comparison more reliable in direct proportion to larger number of cases. To facilitate this comparison, the pretest means for clause length and T-unit length for both sets of essays collected for the present study have been placed in the context of "normative" means previously reported for different age groups, ranging from third graders to skilled adult writers who had published in magazines like Harper's. The context of "developmental norms" for these two syntactic indices appears in Table 4.2.

Insert Table 4.2 about here.

Table 4.2 strongly indicates that as a group the 180 E 306 writers wrote clauses and T-units whose mean length fit well into the developmental continua suggested for both indices. The mean T-unit lengths--15.26 and 16.96, respectively--for both sets of essays are greater than the means Hunt (1965) reported for twelfth graders and smaller than those for the skilled adult writers studied by Cooper, et al. (1980), Hunt (1965), and Faigley (1979a). So too with the mean lengths of the clauses the 180 E 306 students wrote: the average clause lengths of 8.92 words and 10.54 words for the two sets of essays are both greater than the 8.60 words Hunt (1965) reported for twelfth graders and smaller than the means Cooper, et al. (1980), Hunt (1965), and Faigley (1979a) reported for skilled adult writers.

It is also instructive to compare the mean lengths of T-units and clauses written by E 306 students to those reported for freshmen at other universities. As Table 4.2 indicates, both the average T-unit length and the average clause length of the E 306 writers for both essay sets exceeds the means reported for North Dakota freshmen (Failgey, 1979b), New Brunswick freshmen (Stewart, 1978a), and Oklahoma freshmen (Sodowsky and Witte, 1978). The two sets of means for the E 306 students, however, fall on either side of the means reported by Cooper, et al. (1980) for SUNY-Buffalo freshmen. Although some of the differences among the means for the various college freshman populations is assuredly attributable to differences in types of written discourse produced, there is no reason to think that from the standpoint of the development of syntactic complexity, the UT students lag behind their counterparts at other universities. Indeed, if these indices for measuring syntactic complexity are reliable as developmental indices (see O'Donnell, 1976, and Hunt, 1977, for arguments suggesting reliability; and see

Witte, 1979, and Witte and Davis, 1980 and in press, for questions regarding the reliability of these indices), then it is possible to say that from a developmental point of view, the UT students are above average with respect to the development of syntactic complexity.

One additional point needs to be made about the means and the change scores reported in Table 4.1. As was pointed out above, the mean length of clauses and T-units for neither set of E 306 essays increased significantly over the course of one semester. Given the highly significant gain scores reported for these indices for the experimental groups involved in Faigley's (1979 b and c) study of generative rhetoric, Stewart's (1978b) study of sentence combining, and the Miami University trio's (Daiker, et al., 1979; Moreneberg, et al., 1979) study of sentence combining, one might be inclined to wonder whether the E 306 courses had any effect at all on the syntactic abilities of their students. However, it should be pointed out that while significant changes for experimental groups have been reported along these dimensions in a host of sentence-combining experiments and in a few studies of the influence of generative rhetoric, in virtually all instances the change scores along these same dimensions for the respective control groups have indicated either a net loss or a statistically non-significant gain. This suggests that unless students undergo an instructional treatment designed specifically to increase the mean lengths of clauses and T-units, they are not likely to increase the mean lengths of the clauses or T-units they write. Given the previous investigations of

Witte and Davis (1980, in press) on the question of stability of T-unit length, it is even possible to hypothesize that the net losses and the insignificant gains for the various control groups involved in studies designed to promote longer T-units and clauses signal an increased stability along these dimensions in the writing of the control groups.

That the E 306 students failed to realize statistically significant gains along these two dimensions also needs to be placed in the context of two studies which attempted to identify syntactic predictors of writing quality. Neither the study done by Nold and Freedman (1977) nor the study done by Faigley (1979c) indicates that when mean T-unit length and mean clause length are considered as predictors of holistic scores, they fail to account for any significant amount of the variance among the holistic scores. These latter findings are altogether compatible with the results of the T-tests reported in Table 4.1, where statistically significant change is shown to have occurred along the dimension of writing quality for both sets of E 306 essays but not along the dimensions of the two indices of syntactic complexity in question. On the whole issue of predictors of writing quality, we will have much more to say in Chapter 7. For the present, it is sufficient to conclude that these two syntactic indices appear to be useful only as extremely gross indicators of the development of syntactic complexity.

Changes Identified by Scores on Objective Tests. In addition to the measures tied directly to the four essays that each of the 180 E 306 students wrote, we also attempted to measure with a series of objective tests changes along the dimensions of certain writing-related behaviors. The instruments we used were the McGraw-Hill Writing Test, Forms A and B; the McGraw-Hill Reading Test (reading comprehension part only), Forms A and B; and the Daly-Miller Writing Apprehension Test. As with the two sets of writing topics, these instruments were used both as pretests and as posttests. In the case of the two McGraw-Hill tests, half of the students were given Form A as a pretest and half were given Form B; for the posttest the two forms were simply reversed; but the same form of the Daly-Miller Writing Apprehension Test was used as both the pretest and the posttest.

The McGraw-Hill Writing Test reportedly "measures the students' skills in written communication" (Raygor, 1970b, p. 7). Both forms of the test consist of three sections--one devoted to language mechanics (30 items), one devoted to sentence patterns (26 items), and one devoted to paragraph patterns (15 items). These sections of the McGraw-Hill Writing Test were administered sequentially as separate tests and students were allowed 15 minutes to complete each section. Raygor (1970b, p. 7) describes the three sections of the Writing Test as follows:

> Part I - Language Mechanics consists of two short compositions, each of which contains 15 underlined words or phrases. The student must decide whether the underlined material is correct or contains one of the errors listed in the following key:

C error in CAPITALIZATION
P error in PUNCTUATION
G error in GRAMMAR
N NO ERROR

Part II - Sentence Patterns contains a variety of item types. The student must identify sentence types according to the key below:

A Sentence Fragment
B Simple Sentence
C Compound Sentence
D Complex Sentence

He must then examine items which contain four sentences to decide which one is grammatically correct. Finally, he must link together sentences in a short passage by choosing from four given connecting or transition words the one most appropriate in the context of the whole passage.

Part III - Paragraph Patterns contains items which represent paragraphs and require the student to recognize the relationship between the sentences in a paragraph in terms of development of thought. The student's task in some items is (1) to choose the appropriate topic sentence for a paragraph, (2) to choose the appropriate developing sentences for a given topic sentence, and (3) to choose the appropriate

> concluding sentence for a paragraph. In working other items, the student must determine which sentence in a paragraph is out of order. Finally, he is to read through groups of sentences in order to indicate where each group should be divided into paragraphs.

Each of these three sections of the Writing Test generates a separate score, each score having a possible range from zero through the total number of items included in a given section. Collectively, the three sections of the Writing Test yield a total possible raw score of 71.

Although the total possible raw score is the same for both forms of the McGraw-Hill Writing Test, Form A appears to be slightly more difficult than Form B. This difference in level of difficulty is reflected in the means and standard deviations reported by Raygor (1970b, p. 26) for the norming group data. These means and standard deviations appear in Table 4.3.

Insert Table 4.3 about here.

According to Raygor (1970b, p. 26), the "norming group included (1) freshmen (and a few sophomores) in four-year colleges and universities; (2) two-year college students; and (3) 'college-bound' high school juniors and seniors." Depending on the particular form and section of the test, the number of the students in the norming group ranged from 1,171 to 1,204; and the normative data were apparently collected during the 1969 academic year. As Table 4.3 indicates, the difference between the mean scores for the two forms is 1.90 points, suggesting that Form B of the test is somewhat less difficult than Form A, especially as concerns the section on language mechanics. To accommodate for the apparent difference in levels of difficulty between the two forms of the test, the split-halves method of administration noted earlier was employed.

Inasmuch as the "analytic" curriculum underlying many of the E 306 options we examined has a secondary goal the improvement of students' abilities in comprehending written material, we also administered the "Paragraph Comprehension" section of the McGraw-Hill Reading Test, Forms A and B. Raygor (1970a, p. 7) describes this section of the Reading Test in the following way:

> Paragraph Comprehension contains five long reading passages. Each of the five items following each passage measures one of the following five comprehension skills:
>
> 1. Recognition and understanding of the main idea
> 2. Recognition of specific facts and understanding their importance and function
> 3. Recognition and understanding of general scientific principles in both the physical and social sciences

4. Discovery of paragraph organization and structure
5. Critical evaluation of the author's writing: discrimination of fact from opinion, recognition of tone and intent, and judgment of validity of arguments.

> Interspersed among the five longer passages are five short paragraphs with one item each. The latter items measure factual knowledge and are similar to what a student can expect to encounter in an objective test based on a course textbook.

The other two sections of the Reading Test--reading rate and comprehension and skimming and scanning--were not administered simply because of the large chunk of the semester's time that had to be committed to testing in connection with other aspects of the present study; administering the complete Reading Test would have required a minimum of 66 minutes of class time, thus requiring that portions of two class periods in a MWF class would have had to be devoted to the Reading Test alone.

According to Raygor (1970a, p. 30), the "norming group (N = 1,526 for Form A; N = 1,526 for Form B) included approximately equal numbers of (1) freshmen (and a few sophomores) in four-year colleges and universities, (2) two-year college students, and (3) 'college-bound' high school juniors and seniors. As with the norming data for the McGraw-Hill Writing Test, the norming data for the Reading Test were apparently collected in 1969. As with the two forms of the Writing Test, Form A of the paragraph comprehension section of the Reading Test appears to be slightly more difficult than Form B, but not significantly so in this case. This conclusion is based on a comparison of the means and standard deviations presented in Table 4.4 for the two forms of the Paragraph Comprehension section. As Table 4.4 indicates, the means for the two forms of the paragraph comprehension section differ by only 0.64 points along a possible scoring range of zero to 30. Nevertheless, the same split-halves procedure was employed in the administration of the paragraph comprehension section of the Reading Test as was employed in the administration of the two sets of writing topics and in the administration of the Writing Test.

Insert Table 4.4 about here.

One additional point should be made regarding the paragraph comprehension section of the McGraw-Hill Reading Test: it correlates quite highly with the vocabulary, comprehension, and total score scales of Form B of the Nelson-Denny Reading Test, revised edition. The respective Pearson product-moment correlation coefficients reported by Raygor (1970a, p. 32) are as follows: 0.59, 0.54, and 0.61.

While it is not our purpose in the present section of this report to explore the relationships that may or may not exist among these objective measures of writing-related skills (that issue will be taken up in a subsequent chapter), we will assume for the time being that some relationship does exist; and our examination in subsequent portions of the present chapter of the change scores generated with these instruments will reflect that assumption.

The final objective instrument employed in the present study was the Daly-Miller Writing Apprehension Test (WAT). The WAT measures neither writing abilities nor writing-related skills. Rather it attempts to assess the degree of anxiety or apprehension with which writers approach writing tasks. Inasmuch as none of the sections of E 306 we examined for the principal and derivative comparisons specified in Table 2.1 and Table 2.2 appear to have affected writing apprehension, either by increasing it or decreasing it significantly, we will only note the change scores for writing apprehension in the present chapter. However, we do devote an entire chapter later in this report to the relationship between writing apprehension and writing abilities.

The changes observed along the dimensions of these objective tests for the 180 E 306 students, for whom we have already discussed observed changes in two pretest and two posttest essays, are presented in Table 4.5. As Table 4.5 reveals, while no significant changes appeared along the dimensions of the

Insert Table 4.5 about here.

paragraph patterns section of the Writing Test and of the WAT, highly significant changes occurred along the dimensions of the language mechanics section and the sentence patterns sections of the Writing Test and along the dimension of paragraph comprehension section of the Reading Test. The gain scores on the language mechanics and the sentence patterns sections of the Writing Test are shown to be significant beyond the .002 level of confidence, gains which make the changes on the total score of the Writing Test also significant beyond the .001 level of confidence. Given the assumption that some relationship exists between performance on an objective measure of writing skills and actual writing behavior, the higher posttest scores corroborate the changes presented in Table 4.1: the writing abilities, whether measured on essays or on objective tests, of the 180 students appear to have improved over the course of one semester of writing instruction. The gain scores, although less significant from a statistical point of view, on the paragraph comprehension section of the Reading Test are equally impressive since the improvement of reading ability is a secondary goal of E 306 classes. These gains, significant well beyond the acceptable level of .05, are especially impressive when it is remembered that the "synthetic" curriculum sections in the present study made no use of a reader. The data, together with the paired T-tests, presented in Table 4.5 offer additional evidence to suggest that E 306 is an educationally valid course.

4.2 The Five Derivative Comparisons

In Chapter 2, we noted that the research proposal for the present study called for five specific comparisons of E 306 options in which curricular, instructional, and teacher variables were controlled within each comparison. These comparisons, which we labelled the "principal comparisons," are delineated in Table 2.1 in Chapter 2. We also pointed out in Chapter 2 that by removing the controls on the teacher variable and on either a curricular or an instructional variable controlled in the principal comparisons, we could create five "derivative comparisons." Some of these comparisons result in the pairing of instructional or curricular variables not paired in the five principal comparisons, and some merely result in similar comparisons but with a larger number of students because of the removal of the teacher variable.

Each derivative comparison draws on the same data set discussed in the previous section; but each derivative comparison utilized either different combinations of data subsets or different data subsets. The use of different subsets of the E 306 data in the five derivative comparisons will help us identify which curricular and instructional variables had the greatest effect on the changes reported in the previous section. Unlike the five principal comparisons wherein each cell in each comparison is comprised of an equal number (18) of cases, the cells created for the derivative comparisons do not involve equal numbers of cases.

It is to these derivative comparisons that we now turn.

Derivative Comparison 1: All "Analytic-Conventional" Classes Compared with All "Synthetic-Conventional" Classes. This comparison is similar to Principal Comparison 1 in that both test for the effect of curriculum on student writing and writing-related behaviors and skills. The crucial difference between the present comparison and Principal Comparison 1 is that here the number of cases in the "analytic-conventional" classes is 90 rather than 18. The larger N for the present comparison is a direct result of the removal of controls on the teacher variable, thus allowing more classes and more students to be used in the comparison. Because only two of the 20 sections involved in the study were "synthetic-conventional" classes, the N for that component of the present comparison obviously could not be increased.

Before turning to tests of the effect of curriculum on achievement, it is necessary to determine whether the classes on either side of the present comparison registered significant gains during the semester along the twelve dimensions associated with the four essays each student wrote and along the six dimensions associated with the series of objective tests which were administered pre and post. Table 4.6 presents the descriptive data of the pre and post means for the two sets of essays written by the "analytic-conventional" students and the results of T-tests of the observed changes across the period of one semester of writing instruction. Table 4.7 presents the same kinds of information, but for the students in the "synthetic-conventional" classes.

Insert Table 4.6 about here.

Insert Table 4.7 about here.

Some noteworthy differences between the two groups of students appear in Table 4.6 and Table 4.7. As Table 4.7 indicates, students in the "synthetic-conventional" sections showed significant gains along the dimensions of the ratio of T-units with final free modifiers to total T-units and the ratio of words in final free modifiers to total words for the A_2B_2 set of essays; and they showed gains which approached significance along the dimensions of two other syntactic indices. However, the gains for the holistic scores were not statistically significant. Table 4.6, on the other hand, indicates that the "analytic-conventional" students wrote essays on both sets of writing topics which were significantly better at the end of the semester than they were at the beginning. Table 4.6 also indicates that none of the syntactic indices showed gains across the semester that were statistically significant for this group. Although the two tables taken together would seem to suggest that the "analytic-conventional" students outperformed the "syntactic-conventional" students along the dimensions of the two holistic scores--the dimensions which we said earlier must be regarded as the most important--a comparison of the gain scores for the two groups reveals that the change between the pre- and posttest holistic scores for the "synthetic-conventional" students on the A_1B_1 topics was actually greater than for that of the "analytic-conventional" students. An even greater difference, again favoring the "synthetic-conventional" students appears for the A_2B_2 topics. Interestingly enough, the gains along the dimensions of the two holistic scores for the "synthetic-conventional" classes either exceeds or equals those reported for all E 306 students studied (see Table 4.1), gains found to be highly significant. This apparent discrepancy is actually no discrepancy at all, for statistics is the science of probability and high probabilities are difficult to project on the basis of a small number of cases. Thus the probabilty of smaller gains not being attributable to chance alone for a larger number of cases is greater than the probability of larger gains based on a smaller number of cases being attributable to chance. The problem of the small N of cases is one which recurs frequently in the analyses associated with the five principal comparisons presented later in this report as well as in the analyses associated with subsequent derivative comparisons.

The same problem also points to the virtual impossibility of inferring from a comparison of significance levels the relative effectiveness of one or more instructional or curricular components. To make such inferences it is necessary not only to take into account in a systematic way the differences between the pre- and posttest means, but also to accommodate the variance among the scores within the respective populations being studied. An analysis of covariance provides an analytic paradigm which allows all such relevant factors to be accommodated so that reliable inferences can be made regarding the effects of particular treatments on the populations being studied. Hence we employed analyses of covariance as the bases of our comparison of the two curricula taught in a conventional classroom setting. As Table 4.8 reveals, neither the "analytic-conventional" classes nor the "synthetic-conventional" classes can be judged to be more effective than the other in causing changes in posttest A_1B_1 essays. This is not to say that no changes occurred during the span of one semester, but it is to say that one treatment was no more effective than the other in producing those changes.

Insert Table 4.8 about here.

A somewhat different situation arises with respect to the change scores observed for the A_2B_2 topics. As Table 4.9 indicates, there was a statistically significant curricular effect on the pre- to posttest changes in mean number of words, in the ratio of T-units with final free modifiers to total T-units, and in the ratio of words in final free modifiers to total words. By referring back to Table 4.6 and Table 4.7, we can observe that

---

Insert Table 4.9 about here.

---

the "synthetic-conventional" classes produced change scores along these dimensions that were statistically significant and that were greater than those computed for the "analytic-conventional" classes. We can, therefore, infer from the analyses of covariance presented in Table 4.9 that the synthetic curriculum actually effects such changes in student writing.

The "analytic-conventional" and the "synthetic-conventional" classes involved in the present comparison can also be compared along the dimensions of the several objective measures which the students completed. The pre- and posttest means and the results of paired T-tests of the changes across time for the two groups are presented in Table 4.10 and 4.11, respectively.

Insert Table 4.10 about here.

Insert Table 4.11 about here.

As Table 4.10 indicates, the students in the "analytic-conventional" sections achieved significant gains along the dimensions of language mechanics, total writing score, and reading comprehension. The first two of these were statistically significant beyond the .001 level of confidence while the latter was significant beyond the .004 level. The students in the "synthetic-conventional" sections, in contrast, realized no significant gains along the dimensions of reading comprehension--not an unsurprising finding since those students had no reading component in the curriculum they studied--but did realize significant change along all the dimensions of the McGraw-Hill Writing Test. Three of the four change scores, as Table 4.11 shows, reached statistical significance well beyond the .01 level of confidence. One of the change scores, the one for the paragraph patterns portion of the Writing Test, reached statistical significance beyond the .02 level of confidence. However, this change was in a negative direction; that is to say, the students actually regressed along this dimension during the time separating the pretest and the posttest. This regression is somewhat surprising given the amount of attention the students enrolled in the "synthetic" curriculum would have devoted to the study of the paragraph as a microcosm of the whole essay. The highly significant gains along the dimensions of language mechanics and sentence patterns should not, on the other hand, be surprising at all, given the close attention paid to the sentence and to matters related to the sentence in the "synthetic" classes. Obviously, the gains on the total writing score are a function of the significant gains on the two of three sections of the Writing Test. Writing apprehension, as measured by the WAT administered pre and post, changed significantly for neither group.

To determine whether one curriculum could be said to have been more effective than the other in producing the observed change scores, we again ran analyses of covariance to test for the effect of curriculum. As in the previous covariance analyses, the pretest scores, this time for all dimensions reported in Table 4.10 and Table 4.11, were entered into the analyses as covariates. As Table 4.12 shows, a significant curriculum

---

Insert Table 4.12 about here.

---

effect was found only for the performances on the sentence patterns section of the McGraw-Hill Writing Test. When the change scores for this dimension for both groups (see Table 4.10 and Table 4.11) are taken into account, it is possible to conclude that along this one dimension, the "synthetic-conventional" curriculum was more effective than the "analytic-conventional" curriculum.

Given our analyses in the present section, it is possible to highlight what appear to be curriculum effects on the performances of E 306 students. First, neither the "analytic" nor the "synthetic" curriculum when taught in a conventional classroom setting seems to be more effective than the other in producing significant and positive gains in writing quality for either of the two sets of essays. Both curricula appear to be effective in this regard, even though the gain scores for the "synthetic" classes did not reach statistical significance when they were submitted to paired T-tests. Second, although neither curriculum proved to be more effective than the other in producing the change scores observed for the A_1B_1 essays, the "synthetic" curriculum did prove to be more effective in producing changes along the dimensions of mean number of words per essay, the ratio of T-units having final free modifiers to total T-units, and the ratio of words in final free modifiers to total words for the A_2B_2 essay. At this point, these findings suggest to us that the "synthetic" curriculum, in spite of demanding fewer essays of its students, managed to teach "invention" skills in such a way that the students were able to generate significantly more "content" on the A_2B_2 posttest essays than were the students in the "analytic" courses. These findings also suggest that the "synthetic" curriculum taught students to write syntactic structures which have been shown by Nold and Freedman (1977) and Faigley (1979c) to predict a significant amount of the variance among holistic scores. On this matter, we will have much more to say in a subsequent chapter. Finally, the "synthetic" curriculum produced gains along the dimensions of the sentence patterns sections of the McGraw-Hill Writing Test, gains which are probably compatible with the gains along the two syntactic indices which measure the occurrence of final free modifiers.

Derivative Comparison 2: All "Analytic-Conventional" Classes Compared with All "Synthetic-Tutorial" Classes. This comparison is similar to Derivative Comparison 1 in that the "synthetic" and "analytic" curricula are again compared. However, the present comparison controls for neither the teacher variable nor the instructional variable as rigorously as they are controlled in other comparisons (see Table 2.1 and Table 2.2). Differences both in terms of curriculum and in terms of instructional method are involved in this comparison. In this comparison, we studied the same group of "analytic-conventional" students used in the previous comparison, but here we compare their performances in E 306 with those of all students enrolled in the "synthetic-tutorial" classes, regardless of whether those classes used computer-assisted instruction or not. That is to say, in this comparison the data for the "synthetic-laboratory-tutorial" classes (sections 12 and 04) and the "synthetic-tutorial" classes (sections 21 and 03) are pooled.

The change scores for the "analytic-conventional" classes used in this comparison were presented in Table 4.6 and Table 4.10, and their respective significances were discussed above in connection with Derivative Comparison 1. The change scores for the four "synthetic-tutorial" sections used in the present comparison are presented in Table 4.13 and Table 4.14.

Insert Table 4.13 about here.

Insert Table 4.14 about here.

Table 4.13 presents the change scores for the two sets of essays, and Table 4.14 presents the change scores for the objective measures of writing-related behaviors. As Table 4.13 illustrates, the changes along the dimensions of the holistic scores for the two essay sets failed to reach significance at the .05 level of confidence, although the change scores for the A_2B_2 essays did approach statistical significance at the .069 level. It must, however, be noted that the observed positive change of 0.47 for the A_1B_1 essays is only 0.02 of a point short of the gains which were found (see Table 4.1) to be statistically significant for the entire group of 180 E 306 students. Also, the change along the dimensions of holistic scores for the "synthetic-tutorial" students on the A_2B_2 essays actually exceeds the highly significant gains reported for all 180 students (see Table 4.1). Again the problem of the small N appears.

Notice that the change in holistic score for the A_1B_1 essays for the "synthetic-tutorial" classes is 0.19 of a point smaller than that for the "analytic-conventional" classes. For the A_2B_2 essays just the opposite is true, with the holistic change score for the "analytic-conventional" classes being 0.17 of a point smaller than that for the "synthetic-tutorial" classes. Some of the change scores for the "synthetic-tutorial" classes did, however, reach significant levels. As Table 4.13 shows, the "synthetic-tutorial" A_1B_1 posttest essays were significantly longer than their pretest counterparts, as were the A_2B_2 posttest essays, although the level of significance was not as high in the latter case. Significant changes also occurred between the pretest and posttest essays written on the A_2B_2 topics along the dimensions of the ratio of T-units with final free modifiers to total T-units and the ratio of words in final free modifiers to total words. These changes are consistent with those observed for the "synthetic-conventional" classes examined in the previous comparison.

Also consistent with the changes observed previously for the "synthetic-conventional" classes are the changes reported in Table 4.14 for the "synthetic-tutorial" classes. As Table 4.14 indicates, the "synthetic-tutorial" group, although not achieving significant gains along the dimension of the paragraph patterns section of the McGraw-Hill Writing Test, did realize significant gains along the dimensions of language mechanics, sentence patterns, and total score on the Writing Test. These gains are consistent with those reported previously for the "synthetic-conventional" classes. Again, however, the "synthetic" classes in the present comparison failed to realize significant improvement in reading comprehension over the course of the semester while their counterparts in the "analytic-conventional" sections did (see Table 4.10).

Using the data reported in Table 4.6, Table 4.10, Table 4.13, and Table 4.14, we again employed analyses of covariance in order to determine whether either of the two courses represented in Derivative Comparison 2 was more effective than the other. These analyses are presented in Table 4.15, Table 4.16, and Table 4.17.

Insert Table 4.15 about here.

Insert Table 4.16 about here.

Insert Table 4.17 about here.

Table 4.15 indicates that for the A_1B_1 essays neither course can be judged better than the other in effecting significant changes along the dimensions of mean clause length, mean T-unit length, ratio of T-units with final free modifiers to total T-units, ratio of words in final free modifiers to total words, or holistic score. The effect of curriculum and instruction on the holistic scores did, however, approach statistical significance at the .064 level of confidence. The higher mean score for the "analytic-conventional" classes on the A_1B_1 posttest essays was nearly high enough to register as a statistically significant course effect. The posttest mean for the "analytic-conventional" classes was a full three-fourths of a point higher than the posttest mean for the "synthetic-tutorial" classes. The "synthetic-tutorial" course did, however, have a significant effect (significant beyond the .009 level of confidence) on the total number of words written on the A_1B_1 posttest essays. Although the difference of 14.09 words between the posttest means for the two groups may seem too small to be significant at the .009 level, it must be remembered that the pretest scores were used as covariates in this analysis and that the posttest means for the "analytic-conventional" classes represent a net loss from the pretest of 19.80 words (see Table 4.6)

Table 4.16 reports other analyses of covariance, these for the A_2B_2 essay data. As Table 4.16 indicates, neither course had a significantly greater effect than the other on the mean number of words per T-unit, the mean number of words per clause, and the holistic score. Along the other three dimensions, however, significant course effects were found. Table 4.6 and Table 4.13 show statistically significant positive change scores along three dimensions--mean essay length, ratio of T-units with final free modifiers to total T-units, and ratio of words in final free modifiers to total words--for the "synthetic-tutorial" students but not for the "analytic-conventional" students. Thus it would appear that because of the instruction they received, students in the "synthetic-tutorial" classes generated more content for the A_2B_2 essays than did their counterparts in the "analytic-conventional" classes. Furthermore, as the analyses focusing on the use of final free modifiers indicate, the "synthetic-tutorial" students were writing a larger proportion of the syntactic structures shown to be related to writing quality (see Nold and Freedman, 1977; Faigley , 1979c) than were the students in the "analytic-conventional" classes.

Table 4.17 shows that neither course was more effective than the other in producing significant change scores along the several dimensions of the objective measures of writing-related behaviors.

The analyses we have presented in connection with Derivative Comparison 2 suggest that the "synthetic-tutorial" classes provided an attractive alternative for the "analytic-conventional" classes. However, it cannot be determined from these analyses whether the strength of the alternative lies in the difference in curriculum--the "synthetic," Christensen-based curriculum--or the instructional methodology--either tutorial or tutorial with computer-assisted instruction. In an attempt to explore this question, we devised Derivative Comparison 3, to which we now turn our attention.

Derivative Comparison 3: All "Conventional" Classes Compared with All "Tutorial" Classes. This comparison, like the previous two derivative comparisons, involves no control over the teacher variable. In some respects, the present comparison resembles Principal Comparison 4 (see Table 2.1), which was designed to test for the effects of method of instruction on writing performance in the "analytic" curriculum. Unlike Principal Comparison 4, which rigorously controls for the teacher variable and the curriculum variable, the present comparison controls for neither. On either side of the present comparison are included classes representing both curricula. Perhaps the major weakness in the design of the present comparison is that it ignores all curriculum X instruction interactions, thus rendering the findings tentative and inconclusive from the outset.

As with all other comparisons--derivative or principal--reported in the present study, the data subsets used in the present comparison derive from the data set reported in Table 4.1 and Table 4.5 for the 180 students selected from the 20 E 306 sections. For this comparison, the data for all students enrolled in classes taught in conventional lecture-discussion classes (N = 108) were pooled to form one data subset, and the data for all students enrolled in tutorial classes (N = 72) were pooled to form the other data subset.

Table 4.18 and Table 4.19 present the descriptive data for the pretest and posttest dimensions of the two sets of essays for both the "conventional" and the "tutorial" students involved in the present comparison. These two tables also report the results of paired T-tests of the changes across one semester for both groups. Table 4.18 reveals that for the two sets of essays, the "conventional" group showed statistically significant change along three of the twelve dimensions--A_2B_2 essay length, and mean holistic score for both sets of essays. For the "tutorial" group, in contrast, Table 4.19 shows statistically significant positive change along four dimensions, all associated with the A_2B_2 essays. Not only did the "tutorial" group fail to improve their A_1B_1 essays significantly during one semester, but the registered change was only one-fourth of a point on a seven-point scale, about half the change realized by the entire group of 180 students (see Table 4.1) and well less than half the change realized by the "conventional" students. The "tutorial" group realized significant gains during the semester for this set of essays along the dimensions of

---

Insert Table 4.18 about here.

Insert Table 4.19 about here.

---

essay length, ratio of T-units with final free modifiers to total T-units, ratio of words in final free modifiers to total words, and holistic score. The gain in mean essay length of 53.74 words (the equivalent of over three T-units or of over five clauses) is nearly 15 words greater than the change recorded for all 180 students (see Table 4.1) and about 25 words larger than the change for the "conventional" students used in the present comparison. A large gain was also observed for the two syntactic indices focusing on final free modifiers. When the ratios for free modifiers are converted to percentages, we see that the "tutorial" students increased their mean percentage of T-units with final free modifiers by two tenths of one percent, while the "conventional" students increased theirs by only two one-hundredths of one percent. All 180 students increased their percentage by only nine one-hundredths of one percent (see Table 4.1), an increase which failed to reach significance at the .05 level of confidence. Given the previous finding, a comparable impressive change was recorded for the "tutorial" group along the dimension of percentage of words in final free modifiers. The holistic scores on the A_2B_2 essays for the "tutorial" students also changed significantly. The mean change along this dimension was 0.83, which is 0.19 points greater than the gain realized by the "conventional" students and 0.11 points greater than that achieved by all 180 E 306 students.

These data seem to suggest the efficacy of "tutorial" instruction in teaching written discourse of the kind represented by the A_2B_2 essays when the curriculum and teacher variables are left uncontrolled. However, when these data were tested for the effect of the instructional method on the change scores, no significant effects were found, as Table 4.20 and Table 4.21

Insert Table 4.20 about here.

Insert Table 4.21 about here.

---

indicate. With the .05 level of confidence as the criterion level, we note that the effect of the "tutorial" on the production of final free modifiers approached statistical significance at the .085 and the .059 level respectively. When the pretest essays were entered into the analysis as covariates, the lower pretest scores of the "tutorial" students, together with a possible greater variance within the pretest "tutorial" scores, cancelled out the instructional effect.

As some of the gains associated with the two sets of essays were significant for both the "conventional" and the "tutorial" courses, so were some of the gains on the objective tests. As

---

Insert Table 4.22 about here.

Insert Table 4.23 about here.

Table 4.22 and Table 4.23 illustrate, both the "conventional" and the "tutorial" students achieved significant gains on three of the four dimensions of the McGraw-Hill Writing Test. Both groups, failed to achieve significant change scores on the paragraph patterns section of the Writing Test, but both realized statistically significant gains on the sentence patterns section. On the McGraw-Hill Reading Test, the "conventional" group's gain reached statistical significance beyond the .02 level of confidence, while the "tutorial" group's gain failed to reach significance. This latter finding is consistent with those reported earlier for Derivative Comparison 2. In that earlier comparison, the reading comprehension gain score failed to reach significance for the "synthetic-tutorial" students (see Table 4.14), while it did reach significance for the "analytic-conventional" students (see Table 4.10). That the "conventional" classes in both the present comparison and in Derivative Comparison 2 achieved significant gains in reading comprehension and that the "tutorial" classes in neither comparison achieved significant gains suggests that "tutorial" instruction in writing does not enhance reading comprehension to any remarkable degree.

However, when we tested the data for the effect of instructional methodology on reading comprehension, we found that neither method was clearly more effective than the other in this regard. As Table 4.24 illustrates, the effect of instructional method was found to be insignificant when the pretest scores of both groups were used as covariates in analyses of

---

Insert Table 4.24 about here.

---

covariance. Table 4.24 also indicates that neither "conventional" instruction nor "tutorial" instruction is more effective than the other in producing gains along the dimensions of the McGraw-Hill Writing Test. That is to say, neither appears more effective than the other when the classes representing the two curricula in the present study are unevenly, and almost haphazardly, distributed within the two sides of the comparison. However, given the results of the previous two derivative comparisons the apparent equality between the "conventional" and the "tutorial" methods is partly attributable to the contributions of the "tutorial" classes employing the "synthetic" curriculum and partly attributable to the contributions of classes using the "analytic" curriculum.

Derivative Comparison 4: All "Analytic" Classes Compared with All "Synthetic" Classes. Whereas the previous comparison attempted to test for the effects of instructional method on the performances of E 306 students, the present comparison attempts to test for the effects of the two curricula, the "analytic" and the "synthetic." As in Derivative Comparison 3, two important variables in this comparison were left uncontrolled. With the previous comparison, we attempted to test for the effects of instructional methodology, letting the teacher variable and the curriculum variable go uncontrolled. With the present comparison, we attempt to test for the effects of curriculum, letting the teacher variable and the instructional variable go uncontrolled. In some respects the present comparison resembles Principal Comparison 2 and Principal Comparison 5, both of which test for the effects of curricula. However, as Table 2.1 indicates, both principal comparisons control rigorously for both the instructional variable and the teacher variable. Thus in important ways the present comparison may be seen as a prolegomenon to those two principal comparisons. We believed that by removing the controls on the instructional and teacher variables, the resultant larger N would produce a statistically more powerful indication of the relative effectiveness of the two curricula. As in all the derivative comparisons, the data subsets used in the present comparison derived from the data base established for all 180 selected students enrolled in the 20 E 306 classes studied (see Table 4.1 and Table 4.5).

The pretest and posttest data and the results of the paired T-tests of change scores for all "analytic" sections and all "synthetic" sections appear in Table 4.25 and Table 4.26 for the two sets of essays written by both groups of students.

Insert Table 4.25 about here.

Insert Table 4.26 about here.

Table 4.25 reveals that the students enrolled in the "analytic" curriculum achieved significant positive changes only along the dimensions of holistic scores for both sets of essays. This is an important finding, given the principal goals of E 306. The writing of the students, regardless of the instructional methodology, enrolled in the "analytic" curriculum improved over the course of one semester of writing instruction. These change scores for the "analytic" students far exceeded the criterion level of significance at the .05 level. In the case of the A_1B_1 essays, the change was significant beyond the .007 level; and in the case of the A_2B_2 essays, the change was significant beyond the .001 level. These gains by the "analytic" students are nearly identical to those realized by the entire group of 180 E 306 students (see Table 4.1).

Other interesting facts about the performances of the "analytic" students also appear in Table 4.25. For example, the changes recorded for essay length are nearly identical, 18.50 words for the A_1B_1 essays and 18.29 words for the A_2B_2 essays. However, only the change for the A_2B_2 was a positive one: the posttest essays on the A_1B_1 topics were shorter than the pretest essays. Negative change scores were also found along all the syntactic dimensions for the A_2B_2 essays, while positive changes were found along all of these dimensions for the A_1B_1 essays. None of these syntactic changes for either essay set was found to be statistically significant. Thus these findings would seem to suggest that the highly significant positive changes along the dimensions of the two holistic scores are not tied very closely to any of the syntactic features we measured for the "analytic" students.

An examination of Table 4.26, which focuses on the changes recorded for the "synthetic" group, reveals some striking differences between the two groups of students involved in the present comparison, as well as some interesting similarities. For example, the "analytic" students did not significantly increase the lengths of their essays on either set of topics, but the "synthetic" students increased theirs on both beyond the .001 level of confidence. This finding may suggest that the "synthetic" curriculum teaches students to generate more content for a given subject matter than does the "analytic." That is to say, the "synthetic" curriculum may better teach invention skills than does the "analytic" curriculum. A comparison of Table 4.25 and Table 4.26 also reveals that neither group showed any significant change along the dimensions of T-unit or clause length for either set of essays. However, the changes recorded for the two groups along these dimensions were in opposite directions for the two sets of essays. In addition, neither group achieved significant positive change on the A_1B_1 essays along the dimensions associated with final non-restrictive modifiers; but on the A_2B_2 essays the "synthetic" students did achieve highly significant gains (beyond the .001 level of confidence for both indices) along both dimensions. The most important difference between the performances of the two groups lies in the changes in the holistic scores for the two essay sets. Table 4.25 reveals that the "analytic" students realized highly significant gains along this dimension for both sets of essays. Table 4.26 shows that the gains on the A_1B_1 essays for the "synthetic" students failed to reach statistical significance at the .05 level of confidence. This change, however, actually did approach that level of confidence and did, in fact, exceed by 0.05 of a point the highly significant change reported in Table 4.1 for all 180 E 306 students. The gains for the "synthetic" students on the A_2B_2 essays, while significant beyond the .03 level of confidence (a level falling considerably short of the .001 level recorded for the "analytic" students), actually exceeded the 0.72 of a point gain recorded in Table 4.1 for all students and the 0.70 of a point gain reported for the "analytic" students in the present comparison. This apparent discrepancy is actually no discrepancy at all because with an N of 54 (as in the present "synthetic" classes), it is more difficult to compute statistically higher levels of probability than it is with an N of 126 (as in the present "analytic" sections).

Given the respective sample sizes, both groups appear to have fared well along the dimensions that prabably matter the most, the holistic scores assigned to the respective essay sets. Indeed, when we performed the analyses of covariance (using the pretest scores as covariates) to test for the effects of the curricula, we found that neither curriculum was more effective than the other in producing positive changes along the dimensions of the holistic scores. This conclusion is drawn from the analyses summarized in Table 4.27 and Table 4.28. However, given the gain scores reported in Table 4.25 and Table

4.26 and given our discussions of those gain scores, it is

Insert Table 4.27 about here.

Insert Table 4.28 about here.

perhaps better to say that the "analytic" and the "synthetic" curricula are equally effective in producing positive change in writing quality for both types of essays represented by the two essay sets. Other important findings are also indicated by Table 4.27 and Table 4.28. Given the respective changes reported in Table 4.25 and Table 4.26 for essay length on both sets of essays for both groups of students, the conclusion that the "synthetic" curriculum effected the writing of longer essays while the "analytic" curriculum did not is inescapable. This finding may represent evidence that invention is much better taught in the "synthetic" classes than it is in the "analytic" classes. Although we will have much more to say in a subsequent chapter on the relationship between essay length and writing quality, we will note for the present that Nold and Freedman (1977) found that an essay feature called "short" correlated highly with essay quality and that essay length is a reliable predictor of essay quality, a conclusion which is supported by our discussion in Chapter 3 of cohesive ties in E 306 papers rated high or low in quality. The present analyses of variance also point to other important course effects. Given the data and the analyses presented in Table 4.25 and Table 4.26, Table 4.27 and Table 4.28 indicate that while neither the "analytic" nor the "synthetic" curriculum was more effective than the other in promoting changes in mean T-unit and clause length (although the effect of the "analytic" curriculum approached significance), the "synthetic" curriculum seemed to effect highly significant changes in the A_2B_2 essays along the dimensions associated with final free modifiers. As we have repeatedly pointed out, final non-restrictive modifiers have been shown to be important predictors of writing quality (see again Nold and Freedman, 1977; and Faigley, 1979c). However, neither curriculum appeared to have effected comparable changes in the A_1B_1 essays.

Changes for both the "analytic" and the "synthetic" groups were also observed along the dimensions of objective tests of writing-related behaviors. These changes are reported in Table 4.29 and Table 4.30. As these tables illustrate, both groups

---

Insert Table 4.29 about here.

Insert Table 4.30 about here.

---

achieved highly significant changes (beyond the .001 level of confidence) on both the language mechanics section and the total score of the *McGraw-Hill Writing Test*. These tables also show that neither group realized significant changes on the paragraph patterns subtest or on the *Daly-Miller Writing Apprehension Test*. As should be expected, given the attention to sentence-level operations in the "synthetic" curriculum, the "synthetic" students realized significant positive change beyond the .001 level of confidence on the sentence patterns section of the *Writing Test* while the "analytic" students did not. Also as should be expected given the emphasis on reading in the "analytic" curriculum, the "analytic" students achieved gains in reading comprehension which were significant at the .017 level of confidence, whereas the reading scores for the "synthetic" students did not change significantly.

To determine whether a course effect was present in these change scores for the two groups, we again conducted analyses of variance. Again we used the pretest scores for both groups as covariates in the anlayses. The results of these analyses are reported in *Table 4.31*. As *Table 4.31* shows, no significant course effects were found for any of the changes, although those reported for the sentence patterns subtest of the *McGraw-Hill Writing Test* approached signifance at the .08 level of confidence. Given the data and the analyses represented in *Table 4.29* and *Table 4.30*, these findings suggested that course effect

Insert *Table 4.31* about here.

is attributable to the "synthetic" curriculum. The previously reported significant gains for the "analytic" group on the reading comprehension section of the McGraw-Hill Reading Test proved not to be enough larger than those of the "synthetic" students to render higher levels of reading comprehension a function of the "analytic" curriculum.

The data and the analyses we have presented in connection with Derivative Comparison 4 strongly suggest that when the instructional and teacher variables are not controlled, both the "analytic" and the "synthetic" E 306 curricula yield some important and impressive effects on the writing and writing-related performances of the students they serve.

Derivative Comparison 5: All "Analytic-Conventional" Classes Compared with All "Analytic-Tutorial" Classes. This comparison resembles in some ways Derivative Comparison 1 (see Table 2.2). The earlier comparison tested for the effect of curriculum on conventional classroom instruction when the teacher variable was left uncontrolled. The present comparison, in contrast, tests for the effects of the instructional variable within the "analytic" curriculum without controlling for the teacher variable. With two exceptions, the present comparison is identical to Principal Comparison 4 (see Table 2.1); the two differ in that the present comparison does not control for the teacher variable and uses, as a consequence, a larger number of cases on either side of the comparison. As in the previous derivative comparisons, the data subsets here are drawn from the data base reported in Table 4.1 and Table 4.5. In the present comparison, all "analytic-tutorial" sections are pooled in order to form a data subset with 36 cases on one side of the comparison. The data subset for the other side of the comparison consists of the data for all 90 E 306 students who were enrolled in "analytic-conventional" classes. This latter data subset is the same one used previously in Derivative Comparison 1 and Derivative Comparison 2.

Because we have already examined the change scores (see Table 4.6 and Table 4.10) of the "analytic-conventional" students and discussed the significance of those scores, we can focus immediately on the changes achieved by the "analytic-tutorial" students. These changes for the two essay sets are recorded in Table 4.32. As Table 4.32 shows, only the gain on the

Insert Table 4.32 about here.

holistic scores assigned the A_2B_2 essays reached statistical significance. But that gain, while only significant at the .027 level of confidence, exceeds the gain reported for all 180 students involved in the present study (see Table 4.1) by 0.14 of a point; and it exceeds the gain of the "analytic-conventional" students by 0.23 of a point. The holistic scores assigned the A_1B_1 essays, however, are a totally different matter: the change across the semester is almost imperceptible. This latter very small positive change may be partially explained with reference to the unusually high pretest scores on the

A_1B_1 essays, the average of which was 0.56 of a point higher than that for the group of E 306 students as a whole (see Table 4.1).

Some other interesting facts appear in Table 4.32. For example, the A_1B_1 posttest essays were shorter than their pretest counterparts, and the A_2B_2 posttest essays were longer (to a degree approaching statistical significance) than their pretest counterparts. Just the reverse occurred for mean T-unit and mean clause length: whereas the posttest means for both indices were longer for the A_1B_1 essays, they were shorter for the A_2B_2 essays. This finding is altogether compatible with the findings reported for all 180 E 306 students: the entire population gained along these dimensions on the A_1B_1 essays but regressed along them on the A_2B_2 essays. However, the total E 306 population we studied realized gains along the other two syntactic indices on both sets of essays, whereas the "analytic-tutorial" students regressed along those dimensions on the A_1B_1 essays and failed to gain at significant levels on the A_2B_2 essays. Noticeable differences also appear between the "analytic-tutorial" students and the "analytic-conventional" students. Since the means and the changes for the "analytic-conventional" sections were treated previously, we simply refer the reader to Table 4.6 and to the earlier discussions.

To answer the question of the effect of instructional type on student performance in the "analytic-conventional" and the "analytic-tutorial" sections in the present comparison, we again conducted analyses of covariance on the two sets of essay data, allowing the pretest scores for each dimension to function as covariates in the analyses. The results of these analyses are reported in Table 4.33 and Table 4.34. Both tables indicate

Insert Table 4.33 about here.

Insert Table 4.34 about here.

that when the curriculum variable is controlled, one instructional method--whether "conventional" classroom instruction or "tutorial" instruction--seems to have no greater effect on student writing performance than the other. We will have occasion to refer back to this finding in connection with our discussion of Principal Comparison 4, which involves the same comparison but with the teacher variable controlled.

Our analyses of the change scores along the dimensions of the objective tests are not much more revealing at this point. The change scores and the tests of their significance for the "analytic-conventional" sections were reported in Table 4.10 and discussed earlier. The comparable data for the "analytic-tutorial" sections are presented in Table 4.35. This table indicates that the "analytic-tutorial" students realized

Insert Table 4.35 about here.

significant gains along the same dimensions of the McGraw-Hill Writing Test as did the "analytic-conventional" students (see Table 4.10). However, whereas the mean change along the dimension of reading comprehension was significant beyond the .004 level of confidence for the "analytic-conventional" students, the average change for the "analytic-tutorial" students was not at all significant, except in terms of the mean change being very close to zero. As Table 4.10 and Table 4.35 reveal, neither group changed significantly along the dimension of the Daly-Miller Writing Apprehension Test.

As with the previous comparisons, we employed analyses of covariance to determine whether the change scores recorded for the objective tests were the effects of the dependent variable, in this case the instructional method. The results of the analyses are reported in Table 4.36. As Table 4.36 reveals, neither instructional method had a significant effect on the performance of the students examined in the present comparison. However, the differences between the two groups of students did approach statistical significance along two dimensions. With the pretest scores entered as covariates into the analyses of covariance, the difference between the posttest scores suggests that the "analytic-tutorial" courses may have had more of an effect on performance on the language mechanics section than did the "analytic-conventional" courses. The

---

Insert Table 4.36 about here.

---

reverse is suggested for performances on the reading comprehension section of the Reading Test. The difference along that dimension, however, approaches statistical significance only at the .069 level of confidence.

4.3 Chapter 4: A Summary

The preceding five derivative comparisons and the analyses of the performances of all 180 E 306 writers we studied yielded a number of important findings. Paramount among these findings are, of course, those which serve as evidence of the overall effectiveness of E 306.

As we have noted, the overall quality of the narrative-descriptive essays, which were designed to draw on personal experience, improved significantly beyond the .001 level of confidence for the 180-student population we studied. Similarly, the overall quality of the argumentative essays, which adopt a position on an issue and then defend that position, increased substantially across the semester for the entire group of 180 students. This latter gain, which was a good deal larger than the gain registered for the narrative-descriptive essays, was also significant beyond the .001 level of confidence. Our examination of the essay data for all 180 E 306 students also indicates that the students as a group achieved positive gains for two features of the argumentative essays as well. These gains were observed for essay length and mean percentage of words in final free modifiers, two dimensions which previous research has shown to be associated positively with essay quality. Of the one length feature and the four syntactic features we examined on the narrative-descriptive essays, none showed statistically significant change.

We also examined the change scores of all 180 E 306 students for the three subtests of the McGraw-Hill Writing Test, the paragraph comprehension subtest of the McGraw-Hill Reading Test, and the Daly-Miller Writing Apprehension Test. Neither the entire population of 180 students nor any of the various subpopulations we examined registered significant changes on the Writing Apprehension Test, suggesting that neither E 306 nor any instructional or curricular component affects writing apprehension significantly, either by reducing it or by increasing it. For the McGraw-Hill tests, the results were quite different, and much more encouraging, showing that the improvement in writing quality was accompanied by higher scores on objective measures of writing-related skills. The group of 180 students we examined realized gains which were significant beyond the .001 level of confidence on three of the four dimensions of the Writing Test. These dimensions are the language mechanics and the sentence patterns subtests and the total score. Only the changes on the paragraph patterns subtest failed to reach statistical significance at the .05 level of confidence; in fact, the changes along this dimension for the entire 180-student population were in a negative direction. This latter finding, however, may reflect more on the test instrument than the student performances. The entire 180-student group also realized significant gains in reading comprehension. Although those gains did not reach the level of statistical significance observed for the gains on the three dimensions of the Writing Test, they did reach significance beyond the .03 level of confidence. This finding strongly suggests that E 306 not only teaches students to write better but it also teaches them to read better.

Analyses conducted in connection with the five derivative comparisons also yielded some important results. From these analyses, we can conclude that both the "synthetic" and the "analytic" curricula are effective, but in somewhat different ways. The "synthetic" curriculum seems to be more effective in terms of producing longer posttest essays on the argumentative writing topics and in terms of encouraging a greater use of final free modifiers in essays which argue for or against a position on a controversial issue. The "synthetic" curriculum, however, generally appears not to effect increases in the use of final free modifiers on the narrative-descriptive essays, but it does seem to promote longer essays there too. The changes in essay length caused by the "synthetic" curriculum may indicate that that curriculum better teaches invention skills than does the "analytic" curriculum. Neither curriculum appears to be more effective than the other in increasing essay quality, but both appear to promote highly significant positive changes in quality for essays written on both sets of writing topics. This finding is somewhat surprising in light of the fact that the "synthetic" curriculum required its students to write five or six fewer whole essays than did the "analytic" curriculum. It is also somewhat surprising in light of the fact that the "synthetic" curriculum did not specifically teach students to write essays which draw on personal experience, while some attention was paid to expressive discourse in the "analytic" curriculum.

The analyses conducted for the derivative comparisons also suggest some important differences between the two curricula along the dimensions of the objective measures of writing-related skills. Although both curricula appear to promote higher, and statistically significant, posttest scores on the language mechanics subtest and the total score of the McGraw-Hill Writing Test, the "synthetic" curriculum appears to produce positive change on the sentence patterns subtest, whereas the "analytic" curriculum does not. Although the gain scores on reading comprehension subtest of the McGraw-Hill Reading Test for the students in the "synthetic" curriculum were not found to be statistically significant whereas those for the students in the "analytic" curriculum were, analyses of covariance indicated that these differences cannot be attributed to the effect of the "analytic" curriculum.

As regards instructional methods, our analyses are very inconclusive, indicating that both "conventional classroom" instruction and "tutorial" instruction are effective when the curriculum variable is not controlled. When the curriculum variable was controlled, we found no significant effects for the instructional variable on the posttest data for the two essay sets. Our analyses of the posttest data for the objective measures yielded some evidence, although very weak evidence, suggesting that the "tutorial" method improved performance on the language mechanics subtest and that the "conventional" method enhances growth in reading comprehension. While there is very little evidence to indicate that either method better effects changes in writing performance or on scores for objective tests of writing-related skills, the generally higher attrition rate in tutorial classes (see Chapter 1), the frequently higher costs per credit hour, and the larger demands on the teacher's time all suggest that the "tutorial" method of instruction is not an efficient alternative to "conventional" classroom instruction in writing.

CHAPTER 5

THE FIVE PRINCIPAL COMPARISONS

In the preceeding chapter we offered five comparisons which drew in various ways on the complete data set for the 180 students selected for the present study. Those comparisons were not called for in the original research design but seemed to us capable of identifying some of the effects of the two curricula--the "synthetic" and the "analytic"--and some of the effects of the two basic instructional methods--"conventional" classroom instruction and "tutorial" instruction--used in E 306. What those five derivative comparisons showed was that both curricula and both instructional methods are effective, but in somewhat different ways. The five principal comparisons offered in the present chapter will, through the control over the teacher variable, add a degree of specificity not possible in the five derivative comparisons. However, what is gained in specificity is frequently lost in statistical power in the following principal comparisons.

5.1 Review of The Five Principal Comparisons

The five principal comparisons which the present study was initially designed to explore were delineated in Chapter 2. They will be reviewed briefly here.

Principal Comparison 1 is the major comparison. In this comparison, the "analytic" curriculum taught in a conventional classroom setting forms one side of this comparison and the "synthetic" curriculum taught in a "tutorial" setting with the use of computer-assisted instruction forms the other side. These two sides of Principal Comparison 1 represent the two major E 306 options as they were taught during the Fall Semester, 1978. In this comparison both the curriculum and the instruction variables differ across the comparison.

Principal Comparison 2 and Principal Comparison 5 were designed to test for the effects of the two curricula when the instruction variable is controlled. In Principal Comparison 2, the two curricula are compared when the instructional method for both is "conventional" classroom instruction. In Principal Comparison 5, the two curricula are also compared; but in this case, the instructional method used to teach both curricula is "tutorial" instruction, without computer-assisted instruction.

The effects of instructional type is the primary concern of Principal Comparison 4. That comparison was designed to test for the effects of "conventional" classroom instructional and for "tutorial" instruction in teaching the "analytic" curriculum. Ideally, we should have had a similar comparison wherein the effects of the instructional variables could have been tested for the "synthetic" curriculum. Such was not, however, possible because not enough "synthetic" curriculum classes were offered during the Fall Semester of 1978 to accommodate additional comparisons.

Finally, Principal Comparison 3 was designed to test for the effects of the two rhetoric texts, the Rodi syllabus and The Writing Commitment, on learning in "analytic" curriculum classes taught by the "conventional" classroom method of instruction.

With the exception of Principal Comparison 3, it was believed that each of the five principal comparisons would shed some light on the effect of the instructional and curricular variables which, for obvious reasons, could not be controlled in Principal Comparison 1. Each of these comparisons is set out in some detail in Chapter 2 and Table 2.1. Chapter 2 also explains how the teacher variable was controlled for each of these comparisons; and it explains how the data were collected and the sample population selected. As the results of these five comparisons are reported, we shall try to place those results in the context of the performances of the entire group of 180 E 306 students we studied and in the context of the results of the previously reported derivative comparisons. In reporting our analyses and results for the five principal comparisons, we will generally follow the same procedures used in reporting the derivative comparisons.

5.2 The Results of the Five Principal Comparisons

Principal Comparison 1: "Analytic-Conventional" Students (N=18) Compared with "Synthetic-Laboratory/Tutorial" Students (N=18). This comparison is the major comparison of the present study, because it pairs the two major E 306 options as they were taught during the Fall Semester of 1978. It is this comparison which all other comparisons, principal or derivative, were designed to illuminate. In this comparison, as in all principal comparisons, the teacher variable was rigorously controlled. Of the four classes involved in this comparison, all were taught by just two instructors, with each instructor teaching one class on either side of the comparison. Unlike the remaining principal comparisons, the curricular and instructional variables are not rigorously controlled. In this comparison, two curricula and two instructional methods are compared, the "analytic" curriculum taught in a "conventional" classroom setting compared with the "synthetic" curriculum" taught in a tutorial" setting with the aid of computer-assisted instruction.

Table 5.1 and Table 5.2 present the pre- and posttest descriptive data for the two sets of essays for the "analytic-conventional" and the "synthetic - laboratory/tutorial" classes involved in the present comparison. These two tables also report

---

Insert Table 5.1 about here.

---

the results of paired T-tests of the changes across the semester for both groups.

Table 5.1 shows that for the "analytic-conventional" classes, none of the change scores reached significance at the .05 level of confidence, although the gain in mean holistic score for the A_1B_1 essays approached significance at the .069 level. With regard to the A_1B_1 essays, Table 5.1 shows that the "analytic-conventional" students realized negative changes along two of the six dimensions, essay length and clause length. The small loss in mean clause length is not surprising since the "analytic-conventional" students underwent no instruction designed specifically to lengthen clauses. The net average loss of 64.06 words in essay length--an average net loss greater by 44.26 words than that recorded for all "analytic-conventional" students (see Table 4.6)--is important, since the 180 students in all 20 E 306 sections gained, albeit only slightly, as a group along this dimension. The two ratios calculated for the use of final nonrestrictive modifiers in the A_1B_1 are another matter. For these "analytic-conventional" students, the two ratios were considerably larger than those observed for all "analytic-conventional" students (see Table 4.6) and for all 180 students (see Table 4.1). Although the mean holistic score on the A_1B_1 posttest essays of the 18 "analytic-conventional" students represented in Table 5.1 was not significantly higher than that of the pretest essays, the difference between the two was actually quite large. In fact, that average gain of one full point on a seven-point scale is 0.51 of a point larger than the average gain recorded for all 180 students (see Table 4.1) and 0.34 of a point higher than the average gain for all 90 "analytic-conventional" students (see Table 4.6).

The change scores for the A_2B_2 essays are also worth noting. On the A_2B_2 essays there was an average net gain of 22.17 words from the pretest essays to the posttest essays, a gain about ten words larger than that observed for all 90 "analytic-conventional" students (see Table 4.6) and a gain about 17 words smaller than that recorded for all 180 students (see Table 4.1). The average T-unit and clause lengths on the A_2B_2 for these 18 "analytic-conventional" students showed a substantial, although not a statistically significant, increase during the semester. This finding was a little unexpected since the average change for all 90 "analytic-conventional" and for all 180 students along both dimensions was negative (see Table 4.6 and Table 4.1). The mean change scores for the present "analytic-conventional" students along the two dimensions associated with final free modifiers were negative. While the mean change scores for all "analytic-conventional" students (see Table 4.6) along these dimensions were negative, small average group gains were found for all 180 students (see Table 4.1). As was the case for the A_1B_1 essays, the average holistic score on the A_2B_2 essays of the present group of 18 "analytic-conventional" students increased by a full point. This average gain is 0.28 of a point higher than it was for all 180 students (see Table 4.1) and 0.37 of a point higher than for all 90 "analytic-conventional" students (see Table 4.6) used in three derivative comparisons.

Reference to Table 5.2 allows us to compare the average gains of the "analytic-conventional" students represented in Table 5.1 to those of the "synthetic-laboratory/tutorial" students. As Table 5.2 indicates, these 18 "synthetic" students realized statistically significant average gains along two of the twelve dimensions associated with the two essay sets. None of the averages changes for the A_1B_1 essays was significant statistically. Nevertheless, whereas the "analytic-conventional" students in the present comparison had an average net loss of over 64 words in essay length for the A_1B_1 essays, the present "synthetic" students gained on the average 35.5 words, a difference in average change between the two groups of nearly 100 words. Both of the two groups in the present comparison realized similar positive changes in mean T-unit length, and both realized negative changes in mean clause length. The results of these changes are comparable posttest means for the two groups. Whereas the "analytic" students represented in Table 5.1 showed a small average gain for the two ratios for final nonrestrictive modifiers in the A_1B_1 essays, the present group of "synthetic" students recorded an average net loss. Although this average negative change is consistent with the mean negative change for all "synthetic-tutorial" students (see Table 4.13), the average net loss for the present "synthetic" students is greater. For the holistic scores on the A_1B_1 essays of the present group of "synthetic" students, the average change was 0.06 of a point. This mean change is 0.94 of a point smaller than the average gain recorded for the "analytic-conventional" students in the present comparison, 0.43 of a point smaller than the average gain recorded for all 180 students (see Table 4.1), and 0.41 of a point smaller than that of all "synthetic-tutorial" students (See Table 4.13). This latter figure may thus indicate that computer-assisted instruction had an adverse effect on the A_1B_1 posttest essays of the "synthetic-laboratory/tutorial" students.

The "synthetic-laboratory/tutorial" students fared much better on some of the dimensions associated with the A_2B_2 essays. One of these dimensions was not, however, essay length, for which a mean loss of slightly more than two words was recorded. This average difference in length between the A_2B_2 pre- and posttest essays was not only in a negative direction, but it was also over 74 words smaller than that recorded for all "synthetic-tutorial" students (see Table 4.13). It would thus appear that computer-assisted instruction of the type used in the "laboratory" sections had a negative effect on A_2B_2 posttest essay length as well as on A_1B_1 posttest essay quality.

Although these negative effects may be present, there also seem to be some equally impressive positive effects of, apparently, the "computer-assisted" component. Ignoring the insignificant changes in T-unit and clause length on the A_2B_2 essays, we see that positive change was effected along the dimensions of the two ratios associated with final free modifiers and holistic score. First, the two ratios. Converted to percentages, the gains represented by the two ratios--both significant beyond the .002 level of confidence--are quite impressive. The mean percentage of T-units with final nonrestrictive modifiers for the "synthetic-laboratory/tutorial" students rose from 2.9% on the A_2B_2 pretest essays to 7.5% on the posttest essays. This 4.6% average increase is larger by nearly a full percentage point than the average increase for all "synthetic-tutorial" students (see Table 4.13), three and one-half percentage points larger than the average change for all 180 students (see Table 4.1), and three and six-tenths percentage points larger than the average increase for all "analytic-conventional" students (see Table 4.6) and for the "analytic-conventional" students involved in the present comparison. When the ratio of words in final free modifiers is converted to percentages, comparably impressive average gains are observed. Equally impressive average gains were found along the dimension of the holistic score, even though that mean change only approached statistical significance at the .068 level of confidence. The "synthetic-laboratory/tutorial" students improved on the A_2B_2 posttest essays by an average of 1.11 points over the pretest essays. Although the average posttest score of 5.33 is 1.06 points less than the comparable score for the "analytic-conventional" students in the present comparison, the average gain is 0.39 of a point larger than the gain for all 180 students (see Table 4.1), 0.11 of a point larger than the average gain for the "analytic-conventional" students in the present comparison, and 0.31 of a point greater than the average gain for all "synthetic-tutorial" students (see Table 4.13).

To determine whether a course effect caused any of the changes identified in Table 5.1 and Table 5.2, we ran analyses of covariance on the posttest means for both sets of essays. As in the derivative comparisons, the appropriate pretest scores were entered into the analyses as covariates. These analyses for the A_1B_1 and the A_2B_2 essays are presented in Table 5.3 and Table 5.4. As Table 5.3

Insert Table 5.3 about here.

Insert Table 5.4 about here.

indicates, significant course effects were not found for average essay length, T-unit, or clause length on the A_1B_1 posttest essays. However, significant course effects were found for the two dimensions associated with final nonrestrictive modifiers and for holistic score. When the posttest means for the A_1B_1 essays for each course are compared, it becomes obvious that the effects were caused by the "analytic-conventional" course.

Table 5.4 presents the results of the analyses of covariance for the A_2B_2 posttest essay means. As this table illustrates, no significant effects were found for mean clause length, the two ratios associated with final free modifiers, or holistic score. For the latter of the two free modifier indices the F-ratio did, however, approach statistical significance. Comparison of the posttest means of the two groups of students indicates that this high F-ratio favors the "synthetic-laboratory/tutorial" courses. The statistically significant effects for mean essay length and mean T-unit length, on the other hand, indicate that the longer "analytic" posttest essays and T-units are attributable to instruction in the "analytic" classes.

Thus the analyses of covariance we performed on the twelve posttest means for the two E 306 options for both sets of essays suggest that the "analytic" option is the better of the two. No statistically significant course effects for the "synthetic-laboratory/tutorial" classes were found, but five were found for the "analytic-conventional" classes, one of which was for A_1B_1 posttest essay quality.

Not only did the performances of the two student groups differ with respect to actual writing, but two groups of the students also differed with respect to their performances on the objective measures. The pre- and posttest descriptive data for both groups, as well as T-tests of the change scores, appear in Table 5.5 and Table 5.6. As Table 5.5 indicates, only the pretest-to-posttest average change for

---

Insert Table 5.5 about here.

Insert Table 5.6 about here.

---

the language mechanics subtest of the McGraw-Hill Writing Test reached statistical significance for the "analytic-conventional" students. That average gain was 0.60 of a point less than that recorded for all "analytic-conventional" students (see Table 4.10) and 0.88 of a point smaller than the average gain recorded for all 180 E 306 students (see Table 4.5).

Table 5.6, which records the change score averages for the "synthetic-laboratory/tutorial" classes, indicates that significant improvement for that group occurred during the semester along three of the four dimensions associated with the McGraw-Hill Writing Test. Although statistically significant at a higher level than that recorded for the "analytic-conventional" classes, the average change score for the "synthetic" classes on the language mechanics subtest was 0.39 of a point smaller. It is also 1.27 points smaller than the mean gain recorded for all 180 students (see Table 4.5) and 0.61 of a point smaller that the mean gain for all "synthetic-tutorial" students (see Table 4.14). The average gain for the sentence pattern subtest was also statistically significant for the present "synthetic" classes. That mean gain was 0.34 of a point longer than the one for all "synthetic-tutorial" students (see Table 4.14), 1.69 points higher than that for the "analytic-conventional" classes in the present comparison, and one point higher than the mean gain for all 180 students (see Table 4.5). While neither group in the present comparison realized significant gains on the paragraph patterns subtest, the "synthetic" group realized a significant average gain on the total score of the Writing Test, not a surprising finding given the significant gains on the language mechanics and the sentence patterns subtests. This statistically significant mean gain was, however, slightly smaller than the one recorded for all "synthetic-tutorial" students (see Table 4.14), slightly larger than that observed for all 180 students (see Table 4.5), and 1.08 points greater than that of the "analytic-conventional" students in the present comparison.

Although neither group realized statistically significant gains on the paragraph comprehension subtest of the McGraw-Hill Reading Test, the average gain for the synthetic students was 0.81 of a point higher than that of the "analytic" students, suggesting that perhaps computer-assisted instruction positively affects reading comprehension, a suggestion finding some support in analyses presented in connection with other principal comparisons. It should, however, be noted that the mean gain score on the paragraph comprehension subtest for the present "analytic-conventional" students is 0.65 of a point less than that recorded for all 180 students (see Table 4.5) and 1.05 points less than that observed for all "analytic-conventional" students (see Table 4.10). Neither group in the present comparison registered statistically significant changes on the Miller-Daly Writing Apprehension Test.

To determine whether any course effects were indicated for any of these posttest means, we performed analyses of covariance, using the appropriate pretest scores as covariates in the analyses. The results of these analyses appear in Table 5.7. As Table 5.7 indicates, none of the computed F-ratios was statistically significant beyond the .05 level of confidence. The "synthetic" posttest means for the sentence patterns subtest did, however, approach statistical significance at the .08 level of confidence, thus suggesting that the "synthetic" course may with a larger N prove to be effective in producing improvement along that dimension. However, generally speaking, neither the "analytic-conventional" nor the "synthetic-laboratory/tutorial" course appears more effective than the other in producing positive changes measured by these objective tests.

Given the preceeding analyses and presentations of the data for the two groups used in Principal Comparison 1, we must conclude that while the "synthetic-laboratory/tutorial" students performed at least as well as the "analytic-conventional" students on several measures and slightly outperformed them on three, the "analytic-conventional" sutdents clearly outperformed their counterparts along five dimensions, the most important of which was the quality scores assigned to the A_1B_1 essays.

Principal Comparison 2: "Analytic-Conventional" Students (N=18) Compared with "Synthetic-Conventional" Students (N=18). While Principal Comparison 1 allowed for rigorous control of the teacher variable, it did not allow for control of the curriculum and the instruction variables. To have imposed controls on all these variables would have eliminated from the present study the two major E 306 options as they were taught in the Fall of 1978. With the analyses connected with Principal Comparison 1, we were able to determine that in general the "analytic-conventional" classes were more effective than the "synthetic-laboratory/tutorial" classes. However, because neither the curriculum nor the instruction variables could be controlled, the need for additional analyses should be obvious. With the present comparison, we provide additional analyses, focusing on the effects of the two curricula--the "analytic" and the "synthetic"--when both the teacher variable and the instructional variable are controlled. The four classes in the present comparison were taught by two teachers, each teaching one "analytic" class and one "synthetic" class. All four classes were taught in a "conventional" classroom setting. With the controls exerted on the instructional variable, the present comparison thus resembles Derivative Comparison 1; however, in the present comparison with its control over the teacher variable, the same number of cases appear on both sides of the comparison.

The pretest and the posttest essay data for the 18 "analytic-conventional" students and the 18 "synthetic-conventional" students appear in Table 5.8 and Table 5.9. These tables also report the results of paired T-tests of the change scores for the two groups of students.

Insert Table 5.8 about here.

Insert Table 5.9 about here.

Table 5.8 shows that none of the change scores associated with the two sets of essays was statistically significant at the .05 level of confidence for the "analytic-conventional" students. Of the twelve change scores recorded, half were negative and half were positive. For both sets of essays, the mean length of the pretest essays was longer than that of the posttest essays. The average net loss of 16.89 words for the A_1B_1 essays stands in contrast to a very slight average gain for all 180 E 306 students (see Table 4.1), but the loss is approximately three words less than that recorded for all "analytic-conventional" students (see Table 4.6) and about 54 words larger than that for all 180 E 306 students (see Table 4.1). A similar set of circumstances attends the average net loss of 38 words on the A_2B_2 posttest essays for the present 18 "analytic-conventional" students. The posttest mean for the A_2B_2 essays is about 40 words larger than that for all "analytic-conventional" students (see Table 4.6) and about 17 words greater than that for all 180 E 306 students (see Table 4.1). The average net gain for the present "analytic" students is, however, 49 words smaller than the average gain for all "analytic conventional" students (see Table 4.6) and 77 words smaller than that of all 180 students (see Table 4.1).

The syntactic indices also provide the basis for some interesting comparisons. For both sets of essays, the mean T-unit length grew shorter while mean clause length grew longer. These change scores do not appear to be statistically different from those recorded for either all E 306 students (see Table 4.1) or all "analytic-conventional" students (see Table 4.6), even though the directions of the changes among the groups vary considerably. For the two ratios calculated for final nonrestrictive modifiers on both sets of essays, the change scores were not significant. The changes for the present "analytic" students along these dimensions on the A_1B_1 essays were both negative, but they were positive for all "analytic-conventional" students (see Table 4.6); however, the positive changes for this latter group were very small ones. The same holds true for the changes recorded for all 180 E 306 students (see Table 4.1). When ratios are converted to percentages, the positive mean changes of 1.3% and 1% on the two measures of nonrestrictive modifiers should also be compared to the mean changes of all E 306 students and all "analytic-conventional" students. As a group, all E 306 students wrote on the average nearly 1% more of their T-units with final free modifiers on the A_2B_2 posttest essays than on the pretest essays; and they wrote about 0.6 of 1% more words in final free modifiers on the posttest essays than on the pretest essays (see Table 4.1). For all "analytic-conventional" students (see Table 4.6), the reverse was true: those students wrote about 0.4 of 1% fewer T-units with final free modifiers on the posttest essays; and they put about 0.2 of 1% fewer words in final free modifiers.

The recorded changes for the average holistic scores on both essay sets for the present "analytic" students were smaller than those observed for either all E 306 students (see Table 4.1) or all "analytic-conventional" students (see Table 4.6). The 0.33 of a point mean positive change on the A_1B_1 essays for the present 18 students was 0.16 of a point less than the mean gain for all E 306 students and 0.33 of a point smaller than that for all "analytic-conventional" students. Similarly, the 0.44 average positive gain for the 18 students in present comparison was 0.19 of a point and 0.28 of a point smaller than the mean gains for all "analytic-conventional" students (see Table 4.6) and all E 306 students (see Table 4.1), respectively.

For the 18 "synthetic-conventional" students used in the present comparison, statistically significant changes were recorded for none of the six dimensions associated with the A_1B_1 essays and for but three of six dimensions associated with the A_2B_2 essays. Unlike for change scores for the "analytic-conventional" students in the present comparison, the change scores for the "synthetic-conventional" students are all positive.

Although none of the change score means were significant for the A_1B_1 essays, some interesting comparisons can be made. The average increase in word length of 15.39 for the A_1B_1 essays is, for example, about 45 words smaller than the average for all "synthetic-tutorial" students (see Table 4.13); but it is 14 words larger than the average observed for all E 306 students (see Table 4.1) and 32 words larger than the average reported for the "analytic-conventional" students in the present comparison. In this regard, it should be remarked that the mean A_1B_1 posttest essay length is about 60 words greater than that recorded for all E 306 students (see Table 4.1) and about 45 words larger than the one recorded for all "synthetic-tutorial" students (see Table 4.13). Although rather large average gains were realized by the "synthetic-conventional" students on clause and T-unit length, these gains tended merely to increase mean T-unit and clause length to levels comparable to those of all 180 E 306 students (see Table 4.1).

On the two dimensions associated with final free modifiers, these "synthetic" students did quite well. They wrote at least 1% more of their T-units with final free modifiers than did all 180 E 306 students and all "synthetic-tutorial" students. Comparably greater average gains were realized along the dimension of percentage of words in final free modifiers.

The mean gain in holistic score of 0.67 of a point for the "synthetic-conventional" students was 0.61 of a point higher than that observed for the "synthetic-laboratory/tutorial" classes examined in Principal Comparison 1 (see Table 4.38). It was also 0.18 of a point greater than the average gain for all 180 E 306 students (see Table 4.1), and 0.20 of a point greater than the average gain for all "synthetic-tutorial" classes (see Table 4.13).

The A_2B_2 essays for the "synthetic-conventionial" students yielded statistically significant mean changes scores along three dimensions, with a fourth approach-significance at the .065 level of confidence. One of the more important mean changes here is that observed for essay length. While the average change for the A_2B_2 essays for the "synthetic-laboratory/tutorial" classes (see Table 5.2) did not reach statistical significance at the .05 level of confidence, the 117-word average increase for the present group was significant beyond the .001 level of confidence. The average changes for T-unit length and clause length are comparable to those observed for all 180 students (see Table 4.1), all "synthetic-tutorial" students (see Table 4.13), and the "analytic-conventional" students in the present comparison. The significant mean changes associated with the two syntactic indices focusing on the use of free modifiers are both slightly smaller than those recorded for all "synthetic-tutorial" classes (see Table 4.13). However, they are considerably larger than those recorded for all 180 E 306 students (see Table 4.1) and for the "analytic-conventional" students in the present comparison. Although the mean change score for the quality ratings did not reach statistical significance for the "synthetic-conventional" classes, it was identical to the mean change recorded for all 180 E 306 students (see Table 4.1) and 0.28 of a point greater than that for the "analytic-conventional" students in the present comparison. However, it was 0.08 of a point less than that of all "synthetic-tutorial" students (see Table 4.13).

To determine whether one or the other curriculum could account for the differences between the posttest means on the two sets of essays, we again conducted analyses of covariance, using the pretest scores as covariates. As Table 5.10 and Table 5.11 indicate, only one curricular effect appeared. And it is clear from a comparison of the mean lengths of the A_2B_2 posttest essays for the "analytic-conventional" and "synthetic-conventional" students that the "synthetic" curriculum effected longer essays than the "analytic" curriculum did. Both curricula taught in a conventional

Insert Table 5.10 about here.

Insert Table 5.11 about here.

classroom setting appear to be equally effective in promoting change along the other eleven dimensions associated with the two essay sets.

We also compared the "analytic-conventional" classes and the "synthetic-conventional" classes along the five dimensions associated with writing-related skills and along the changes recorded for the Daly-Miller Writing Apprehension Test. These changes, together with paired T-tests of their significance, appear in Table 5.12 and Table 5.13. As Table 5.12 shows, the "analytic-conventional" students realized statistically significant

Insert Table 5.12 about here.

Insert Table 5.13 about here.

average gains on the language mechanics and the sentence patterns subtests of the McGraw-Hill Writing Test. These gains, in turn, produced a significant gain on the total score of the Writing Test. The average gain on the language mechanics subtest was only slightly larger than that observed for all 180 students in the present study (see Table 4.5) but about one-half point greater than that for all "analytic-conventional" students (see Table 4.10). For the sentence patterns subtest, the average gain was 1.06 points and 1.63 points greater than those recorded for all 180 students and for all "analytic-conventional" students, respectively. Not unexpectedly, comparably larger mean gains occurred for the total score on the Writing Test. The average gain of 5.31 points is 2.21 points larger than that recorded for all "analytic-conventional" students and 1.59 points greater than the one recorded for all 180 E 306 students. Although the present "analytic-conventional" students did not realize statistically significant average gains on the paragraph comprehension component of the McGraw-Hill Reading Test, that gain approached statistical significance at the .074 level of confidence. Although in itself not significant at or beyond the .05 level of confidence, the average gain in reading comprehension of 1.76 points exceeds by 0.58 of a point the average gain recorded in Table 4.10 for all "analytic-conventional" students and by 0.98 of a point the average gain observed for all 180 E 306 students (see Table 4.5).

When Table 5.13 and Table 5.12 are compared some interesting differences, as well as some similarities, appear between the two groups in the present comparison. For example, the mean change for the 18 "synthetic" students on the language mechanics subtest is 3.94 points, a positive change 0.56 of a point larger than the one recorded for the 18 "analytic" students. A similar difference between the mean changes on the sentence patterns subtest can be observed, with the mean change score for the "synthetic" students 0.68 of a point larger than the one for the 18 "analytic" students. On the paragraph patterns subtest the "analytic" students realized only a very slight gain of 0.06 of a point, whereas the "synthetic" students moved in a negative direction by 1.31 points on the average. The difference between the total score mean changes, both of which were positive and highly significant, is only 0.12 of a point. On the paragraph comprehension subtest, the difference between the two groups' performances are especially important. Whereas the "analytic" students realized a positive mean change that approached statistical significance, the "synthetic" students on the average registered a negative change. When the positive change and the negative change for the two groups are summed, the difference between the change score means is seen to be 2.17 points. Neither of the two groups realized significant change on the Daly-Miller Writing Apprehension Test.

Analyses of covariance were also performed to compare the objective test posttest mean scores recorded for the "analytic-conventional" and the "synthetic-conventional" classes. In those analyses, the pretest scores of each group were entered as covariates. With these analyses we sought to determine whether the change scores recorded in Table 5.12 and Table 5.13 could be attributed to the effect of either the "synthetic" or the "analytic" curriculum. As Table 5.14 indicates, no course effects were observed, in spite of the differences between the mean change scores of the two groups. That is to say, neither the "analytic" curriculum nor the "synthetic" curriculum was shown to be more effective than the other in producing changes along the several dimensions of the objective measures of writing-related skills or for the Daly-Miller Writing Apprehension Test.

Insert Table 5.14 about here.

Principal Comparison 2 was designed to determine whether the five significant course effects found for the "analytic-conventional" E 306 option studied in Principal Comparison 1 were attributable to the curriculum. Given the present analyses of covariance for the A_1B_1 and the A_2B_2 essays and the objective measures for the "analytic-conventional" and the "synthetic-conventional" classes, we cannot conclude that the course effects determined in Principal Comparison 1 were in fact caused by the "analytic" curriculum. In the present comparison the only statistically significant curriculum effect that was found was for the average length of the A_2B_2 essays. That effect was attributable not to the "analytic" curriculum but to the "synthetic" curriculum. It should be noted that a course effect was found for the same dimension in Principal Comparison 1; but that effect, significant beyond the .003 level of confidence, was caused by the "analytic" course, not the "synthetic" course. Tentatively, we can thus conclude that with the one exception noted above, the two curricula are equally effective or ineffective when they are taught in a conventional classroom setting. Our analyses also suggest that either the "tutorial" component or the "computer-assisted laboratory" component of the "synthetic-laboratory/tutorial" courses examined in Principal Comparison 1 is instructionally less effective than "conventional" classroom instruction in writing. These latter questions will be explored in a subsequent comparison.

Principal Comparison 3: "Analytic-Conventional" Students (N=18) Compared with "Analytic-Syllabus-Conventional" Students (N=18). From the analyses conducted for Principal Comparison 1 and Principal Comparison 2, it cannot be concluded whether the effects of the "analytic" courses are attributable to the Rodi-syllabus. This is the question addressed in Principal Comparison 3. As in the other principal comparisons, the teacher variable in Principal Comparison 3 is rigorously controlled. Each of the two teachers involved taught one section which used only the Rodi-syllabus, and each taught one section which used Adelstein and Pival's The Writing Commitment (1976) together with the Rodi-syllabus. Even though Principal Comparison 3 was designed to test for the effect of the rhetoric text, the text variable could not be as rigorously controlled as, ideally, it might have been. Ideally, different rhetoric texts would have paralleled different curricula, as they did in Principal Comparison 2. The four classes used in the present comparison were, however, all taught the "analytic" curriculum, and both teachers organized their two classes around the chapters in the Rodi-syllabus. In two of the four classes, one taught by each of the two teachers, the "analytic" curriculum as presented in the Rodi-syllabus was supplemented by rhetoric-related topics in The Writing Commitment. All other instructional materials--e.g., handbook and reader--were the same for all four classes.

The pretest and posttest essay data for the 18 "analytic-conventional" students and the 18 "analytic-syllabus-conventional" students used in the present comparison are summarized in Table 5.15 and Table 5.16. These tables also report the results of paired

---

Insert Table 5.15 about here.

Insert Table 5.16 about here.

T-tests of the changes for the two sets of essays for both groups. Table 5.15 indicates that statistically significant changes along four dimensions occurred for the group that used The Writing Commitment, while a fifth approached significance at the .07 level of confidence. Of the four changes which were found significant and the fifth which approached significance, all but one were negative gains. That is to say, the pretest essay means for those five dimensions were, with one exception, all larger than the posttest means. Indeed, of the changes reported in Table 5.15, only four of the twelve represented positive gains--the changes in mean word length and mean holistic score for both essay sets. The change in mean word length of the A_1B_1 essays was over 12 words greater than that observed for all 180 E 306 students (see Table 4.1) and over 33 words greater than that for all "analytic-conventional" students (see Table 4.6). The mean change on the A_2B_2 essays was about the same as that recorded for "analytic-conventional" students, but it was about 30 words smaller than that for all 180 students (see Table 4.1). It is also noteworthy that the pre- and posttest means for both essay sets are considerably smaller than those observed both for all 180 students and for all "analytic-conventional" students.

Table 5.15 also indicates that while a rather large mean positive gain was found for the A_1B_1 holistic scores, a rather small one was found for the A_2B_2 holistic score. The average gain of 1.44 points for the A_1B_1 essays is 0.95 of a point larger than that recorded for all 180 students (see Table 4.1) and 0.78 of a point larger than that of all "analytic-conventional" students (see Table 4.6). The smaller gains for both of these two latter groups, it will be remembered, were highly significant. The average gain for the A_2B_2 holistic scores was much smaller than the gain on A_1B_1 essays. The registered mean gain of 0.39 of a point is 0.33 and 0.24 of a point smaller than those recorded for the 180 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6), respectively. These gain scores would seem to suggest that these two "analytic-conventional" classes in the present comparison better learned to write narrative and descriptive essays based on personal experience than they did to argue objectively for a particular position on a controversial issue. However, an examination of the pretest A_2B_2 holistic scores and the posttest scores suggests that, in comparison to all 180 E 306 writers studied (see Table 4.1), these 18 students were already fairly competent writers at the beginning of the semester. The pretest A_2B_2 average holistic score was 4.89, 0.42 of a point larger than the average pretest score for all 180 students. The 0.42 of a point difference represents about 58% of the total change recorded for all 180 students. In contrast, on the A_1B_1 essays, with a pretest mean of only 4.06 compared with 4.52 for all 180 students (see Table 4.1), these students had a long way to go; and they did.

The mean changes on the syntactic measures for the present group of "analytic-conventional" students were all negative; and three of these eight negative changes were statistically significant well beyond the .05 level of confidence. These changes are somewhat puzzling. While not altogether inconsistent with the changes recorded for all "analytic-conventional" students (see Table 4.6), these changes do suggest that the students were writing shorter and simpler sentences at the end of the semester than they were at the beginning. These negative syntactic gains may thus suggest that one or both of the two teachers involved in the present comparison strongly favored a "simple," direct writing style over a more complex, elaborate style. If the syntactically simpler style were systematically taught, it was taught well: the mean holistic scores for both sets of posttest essays exceed those recorded for all 180 E 306 students (see Table 4.1). It should also be pointed out in this connection that either statistically insignificant gains or negative gains frequently appear in reports of experimental composition research for "control" groups. Typically, these "control" groups were taught an "analytic" curriculum.

The data summarized in Table 5.16 for the two "analytic-syllabus-conventional" classes indicate that while half of the twelve mean change scores represented positive change, none was statistically significant. As was observed of the mean essay lengths summarized in Table 5.15, the average length of the essays written by the "syllabus" students is generally smaller than that reported for all 180 E 306 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). It is unclear to us why these smaller means occurred for both groups on the present comparison on both the pretest and posttest essays in both essay sets.

Although smaller average essay lengths were observed for both groups in the present comparison, some noticeable differences between the two groups also were found, especially for the syntactic indices. Whereas the "analytic-conventional" group in the present comparison lost considerable ground on the four A_1B_1 syntactic indices, only the average T-unit length of the "syllabus" students decreased over the semester. That decrease, however, served to bring the posttest means for that group on that index more in line with the posttest means noted for all 180 students (see Table 4.1). The small gain in average length on the A_1B_1 essays, although about twice as large as that observed for all 180 students (see Table 4.1), resulted in posttest mean clause lengths comparable to those for all students but still 1.31 words longer than those for the "analytic-conventional" students used in the present comparison. The positive mean gains recorded for the "syllabus" students for the two ratios associated with final free modifiers indicate a heavier reliance on free modifiers in the A_1B_1 essays of the "syllabus" students than in the A_1B_1 essays of the "analytic-conventional" students used in the present comparison, of all "analytic-conventional" students (see Table 4.6), and of all 180 students (see Table 4.1).

The average pretest-posttest change for the holistic scores on the A_1B_1 essays for the "syllabus" students was 0.39 of a point, which was 0.27 of a point less than that for all "analytic-conventional" students (see Table 4.6), 0.10 of a point less than that for all 180 students (see Table 4.1) and 1.05 points less than that of the "analytic-conventional" students in the present comparison. Not only was the average gain for the holistic scores less than it was for the students in those other three groups, but so was the mean score on the posttest A_1B_1 essays. In fact, the mean holistic score on the A_1B_1 posttest essays for the "syllabus" students was over three-fourths of point smaller than that of the "analytic-conventional" students in the present comparison.

As with the A_2B_2 change scores for the "analytic-conventional" students in the present comparison, two of those change scores for the "syllabus" students were positive and four were negative. Negative changes were recorded for both groups on all four syntactic indices, while positive changes were found for average essay length and mean holistic score. For the "syllabus" students, however, the positive changes were larger and the negative changes for mean T-unit and mean clause were smaller. The negative changes observed for the two ratios associated with free modifiers were larger than those recorded for the "analytic-conventional" students in the present comparison, as were the positive changes in average essay length and mean holistic score. All of the posttest means recorded for these "syllabus" students are comparable to those observed for all "analytic-conventional" students (see Table 4.6). With the exception of the two ratios associated with final free modifiers, the posttest means for the "syllabus" students are also compatible with those observed for all 180 students (see Table 4.1). Whereas the "syllabus" students used, on the average, final free modifiers in 3% of the T-units they wrote, the entire sample of 180 students used such modifiers, on the average, in 4.8% of their T-units. Similarly, whereas 1.6% of all words in the A_2B_2 essays of the "syllabus" students appeared in final nonrestrictive modifiers, 2.7% of all words used by the 180 students appeared in final free modifiers.

To determine the effects of the "rhetoric text" on the posttest means displayed in Table 5.15 and Table 5.16, we conducted analyses of covariance on each, using in each case the pretest scores as covariates. The results of these analyses are presented in Table 5.17 and Table 5.18.

Insert Table 5.17 about here.

Insert Table 5.18 about here.

As Table 5.17 indicates, a statistically significant effect was found only for mean clause length on the A_1B_1 essays. That effect, which favors the "syllabus" classes, is largely attributable to the fact that on this dimension the "analytic-conventional" students regressed slightly during the semester while the "syllabus" students gained slightly. Although statistically significant effects were found for none of the remaining five dimensions, the computed F-ratios for four of these dimensions approached significance. Three of these F-ratios indicate ever so tentatively that the Rodi syllabus had a positive effect on mean T-unit length, mean ratio of T-units with final free modifiers to total T-units, and mean ratio of words in final free modifiers to total words. The fourth

F-ratio suggests, again only in a tentative way, that the use of The Writing Commitment in conjunction with the Rodi syllabus had some bearing on the holistic scores of the A_1B_1 essays. Thus we would tentatively suggest that the Rodi syllabus used without The Writing Commitment effected certain syntactic changes in the type of discourse represented by the A_1B_1 writing topics, while its use in conjunction with The Writing Commitment had some effect on the overall quality of such written discourse.

The analyses of covariance for the dimensions analyzed for the A_2B_2 essays are represented in Table 5.18. These analyses yielded results much less inconclusive than those derived from the analyses of the A_1B_1 essays. Clearly, neither the Rodi syllabus used alone nor The Writing Commitment used in conjunction with the Rodi syllabus had a greater effect on mean essay length, the two ratios associated with final nonrestrictive modifiers, or holistic score. Statistically significant effects were, however, found for mean T-unit length and mean clause length. These effects, significant beyond the .009 level of confidence, clearly favor the Rodi-syllabus used without The Writing Commitment.

We also compared the performances of the two groups on the objective measures of writing-related skills and of writing apprehension. The pre- and posttest means for the two groups, together with paired T-tests of the change scores for both groups, are presented in Table 5.19 and Table 5.20. Table 5.19 reveals

---

Insert Table 5.19 about here.

Insert Table 5.20 about here.

---

that none of the changes for the "analytic-conventional" students, the students exposed to both The Writing Commitment and the Rodi syllabus, reached statistical significance at the .05 level of confidence. The positive gains in reading comprehension did, however, approach significance at the .054 level. The average change score on this measure of 2.27 points is, it should be noted, 1.09 points higher than the average change recorded by all "analytic-conventional" students (see Table 4.10), a change significant beyond the .004 level of confidence. This average change of 2.27 points was also higher by 1.49 points than the average change noted for all 180 students, change which was determined to be statistically significant beyond the .03 level of confidnece. The change recorded for writing apprehension was neither large nor significant, a finding compatible with those for every other group of students investigated in the present study. Of the four sets of scores on the McGraw-Hill Writing Test, two were positive, one was negative, and one represented no change at all; but none of the change scores was statistically significant. The mean change of 1.71 points observed for the language mechanics subtest was smaller by 1.41 points than the average change noted for all 180 students (see Table 4.5) and smaller by 1.03 points than that for all "analytic-conventional" students (see Table 4.10). The average change, a negative one, on the sentence patterns subtest was 0.54 of a point less than that of all "analytic-conventional" students (see

Table 4.10) and 1.11 points less than that of all 180 students (see Table 4.5). For the present "analytic-conventional" students, the average change on the paragraph patterns subtest was zero. The total writing score mean gain was 1.41 for the present students. This figure is 1.69 points lower than the one for all "analytic-conventional" students (see Table 4.10) and 2.31 points lower than the one for all 180 students (see Table 4.5).

Table 5.20 reveals that substantially larger changes on two dimensions of the Writing Test were found for the "syllabus" students. The mean change for the "syllabus" students on the language mechanics subtest was 4.29 points, 2.58 points higher than that observed for the "analytic-conventional" students in the present comparison. As Table 5.20 also indicates, this mean change for the "syllabus" students was significant beyond the .001 level of confidence. This highly significant and large positive gain on the language mechanics subtest produced a statistically significant average change in total score. Significant beyond the .02 level of confidence, this mean change for the "syllabus" students was 0.87 of a point greater than that for all 180 students (see Table 4.5), 1.49 points larger than that for all "analytic-conventional" students (see Table 4.10), and 3.18 points greater than that for the "analytic-conventional" students in the present comparison. Although the average gain of 1.11 points on the reading comprehension measure was not statistically significant, that gain was actually larger than the statistically significant average gain recorded for all 180 students (see Table 4.5) and only slightly smaller than the gain for "analytic-conventional" students (see Table 4.10). The average gain in reading comprehension for the "syllabus" students was, however, 1.16 points smaller than that observed for the "analytic-conventional" students in the present comparison.

To determine the effects of the "rhetoric text" on student performances on these objective measures, we ran analyses of covariance on the posttest scores, using the corresponding pretest scores as covariates. The results of these analyses appear in Table 5.21. This table indicates that neither the Rodi syllabus used alone nor The Writing Commitment used in conjunction with the Rodi syllabus had a greater effect on the posttest means. Both appear to be equally effective in some ways and equally ineffective in others.

Insert Table 5.21 about here.

The analyses we performed in connection with Principal Comparison 3 allow a few speculations. First, while neither "rhetoric" nor combination of "rhetorics" seems to be more effective in producing qualitative changes in the A_2B_2 essays, the argumentative essays, the Rodi syllabus itself seems to have an adverse effect on the quality of the A_1B_1 essays. On the other hand, it may have effected longer T-units and clauses and a greater use of free modifiers. At any rate, the question of the relative effectiveness of the Rodi syllabus and The Writing Commitment is a moot one because one has been dropped altogether from the "analytic" curriculum and the other's use is limited to that of a training manual for new teaching assistants.

Principal Comparison 4: "Analytic-Conventional" Students (N=18) Compared with "Analytic-Tutorial" Students (N=18). This comparison was designed to test for the effect of the instructional method on the writing and writing-related performances of students studying the "analytic" curriculum. The students taught in a "conventional" classroom setting and the students taught with the "tutorial" method used the same set of instructional materials and followed the same course outline, which was determined in large part by the Rodi syllabus.

The pretest and the posttest essay data for the 18 "analytic-conventional" students and the 18 "analytic-tutorial" students used in the present comparison are summarized in Table 5.22 and and Table 5.23. These tables also present the results of paired T-tests of the mean change scores observed for the two groups along the twelve dimensions associated with the A_1B_1 essays and the A_2B_2 essays.

Insert Table 5.22 about here.

Insert Table 5.23 about here.

Table 5.22 shows that mean length of the A_1B_1 posttest essays was over 19 words shorter than that of the pretest essays for the 18 "analytic-conventional" students. This negative change, although not statistically significant, resulted in A_1B_1 posttest essays which were, on the average, 42.87 words shorter than the A_1B_1 posttest essays of all 180 E 306 students (see Table 4.1) and 34.52 words shorter than the A_1B_1 posttest essays of all "analytic-conventional" students (see Table 4.6). The present "analytic-conventional" students realized an average positive gain in the number of words on the A_2B_2 essays; and that gain, although small, was comparable to the average gains recorded for all "analytic-conventional" students (see Table 4.6). However, the A_2B_2 posttest essays were shorter, on the average, than those of either of the other two groups. The average word length on the A_1B_1 posttest essays of the present "analytic-conventional" students was 390 words. Thus the A_1B_1 posttest essays for these 18 students were on the average 87.4 words shorter than those of all 180 students (see Table 4.1) and 64.4 words shorter than those of all "analytic-conventional" students (see Table 4.6).

Of the mean change scores along the eight syntactic dimensions for these 18 "analytic-conventional" students on the two essay sets, only one was significant beyond the .05 level of confidence. That change score--the one for mean T-unit length on the A_1B_1 essays --suggests that the growth along that dimension allowed the students to achieve mean T-unit lengths on their posttest essays that were nearly identical to those recorded for all 180 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). The same observation holds true for the mean change in clause length, even though that change was not statistically significant. The positive changes for this group along the two dimensions associated with final nonrestrictive modifiers are another matter. The mean changes on these indices are substantially larger than those recorded for all 180 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). The result of these larger change scores were A_1B_1 posttest essays which had at least 1.3% more T-units with final free modifiers than did the comparable essays of all 180 students (see Table 4.1) and of all "analytic-conventional" students (see Table 4.6). Similarly, the posttest A_1B_1 essays of these 18 students used about 1% more of their total words in final free modifiers than did the A_1B_1 posttest essays in the other two groups. The syntactic changes recorded for these 18 students on the A_2B_2 essays were mixed changes, two positive and two negative. Although the positive change in mean T-unit length was insignificant statistically, the A_2B_2 posttest mean for these "analytic-conventional" students was more than a half a word larger than the comparable means for all 180 E 306 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). The mean changes recorded for the final free modifiers were, however, both negative. As a result, a lower percentage of T-units with final free modifiers appeared in the A_2B_2 posttest essays of these 18 students than appeared in the corresponding posttest essays of all 180 students (see Table 4.1) and of all "analytic-conventional" students (see Table 4.6). The same can also be said of the mean percentage of total words used in final free modifiers in the A_2B_2 posttest essays of the 18 "analytic-conventional" students.

As Table 5.22 shows, these 18 students also realized positive gains on the holistic scores for both essay sets. However, neither of these average gains was found to be statistically significant beyond the .05 level of confidence. In fact, the average pretest-posttest gain on the A_1B_1 essays was quite small, the improvemnt in writing quality registering only 0.11 of a point on a seven-point scale. This almost negligible gain resulted in a mean holistic score on the A_1B_1 posttest essays that was 0.41 and 0.58 of a point less than those recorded for all 180 E 306 students (see Table 4.1) and all "analytic-conventional" students (see Table 4.6), respectively. Although the average gain on the A_2B_2 essays was not statistically significant, that gain equalled that of all 180 students (see Table 4.1) and was somewhat larger than that for all "analytic-conventional" students (see Table 4.6). However, the posttest mean holistic score was 0.64 and 0.82 of a point less than those recorded for the other two groups (see Table 4.1 and Table 4.6). In this connection, it should be pointed out that the average pretest holistic score for this group of 18 students was also substantially lower than those of the other two groups.

As with the mean change scores of the 18 "analytic-conventional" students in the present comparison, only one of the mean change scores for the 18 "analytic-tutorial" students reached statistical significance beyond the .05 level of confidence. While eight of the mean change scores for the former group were positive, only six for the "tutorial" students were; and four of those were for the A_2B_2 essays.

Table 5.23 shows that the 18 "tutorial" students wrote, on the average, posttest essays shorter than the pretest essays on the A_1B_1 topics, while the reverse was true for the essays written on the A_2B_2 topics. Although an average negative change of 37.61 words in essay length between the pretest and posttest A_1B_1 essays was observed for the "analytic-tutorial" students, the mean length of the posttest essays was about the same as that recorded for all 180 E 306 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). However, the mean length of the A_1B_1 essays for these "tutorial" students was about 45 words larger than that recorded for the 18 "analytic-conventional" students in the present comparison. In contrast to that found for the A_1B_1 essays, the mean change in essay length for the "tutorial" students' A_2B_2 essays was positive. The average posttest gain over the pretest means for essay length was 21.67 words. This mean change falls in between those recorded for all 180 E 306 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6); and the posttest essay means are comparable to those observed for both of the other two groups. However, the average length of the A_2B_2 posttest essays of the "tutorial" students was nearly 79 words larger than that of the 18 "analytic-conventional" students in the present comparison.

There were also some marked differences between the two groups in the present comparison along the dimensions of the eight syntactic indices. Whereas all of the recorded syntactic changes on the A_1B_1 essays were positive for the 18 "conventional" students, only two were positive for the "tutorial" students. Both groups showed positive average gains in T-unit and clause length; and although the posttest means on both measures are slightly larger for the 18 "conventional" students, those means are roughly equivalent to the ones recorded for the "tutorial" students. As were the posttest means along these two dimensions for the 18 "conventional" students, those for the 18 "tutorial" students compare well with the A_1B_1 posttest means recorded for all 180 E 306 students (see Table 4.1) and for all "analytic-conventional" students (see Table 4.6). Along these two measures for the A_2B_2 essays, the 18 "tutorial" students regressed somewhat, producing posttest means that were smaller than those observed for the 18 "conventional" students in the present comparison. Although the average changes in T-unit and clause length for the present "tutorial" students were negative, the resultant posttest means are comparable to those produced by all 180 E 306 students (see Table 4.1) and all "analytic-conventional" students (~~see~~ Table 4.6).

Differences between the two groups in the present comparison also occurred along the four dimensions associated with the use of final free modifiers. For the 18 "conventional" students the average changes on the A_1B_1 essays were positive, while on the A_2B_2 essays they were negative. Just the reverse was true for the "tutorial" students: they registered negative gains on the A_1B_1 essays and positive ones on the A_2B_2 essays. Although negative mean changes were recorded for these indices on the A_1B_1 essays of the 18 "tutorial" students, the posttest means are still somewhat higher than those recorded for all 180 students (see Table 4.1), for all "analytic-conventional" students (see Table 4.6), and for the 18 "analytic-conventional" students in the present comparison. In contrast to these negative changes of the "tutorial" students on the A_1B_1 essays stand positive mean changes on the A_2B_2 essays. Although these average changes were not statistically significant, they did produce A_2B_2 posttest means which were substantially higher than those recorded for all 180 students (see Table 4.1), all "analytic-conventional" students (see Table 4.6), and the 18 "analytic-conventional" students in the present comparison. In terms of the percentage of T-units with final free modifiers, these differences indicate that, on the average, the present "analytic-tutorial" students wrote a higher percentage of T-units with final free modifiers than did the present "analytic-conventional" students, the 180 E 306 students, and all "analytic-conventional" students respectively.

The A_1B_1 and the A_2B_2 average holistic scores also differed for the two groups in the present comparison. The 18 "analytic-conventional" students realized a small mean change between the pretest and posttest A_1B_1 essays but still produced posttest essays with an average score smaller than those recorded for all 180 students and for all "analytic-conventional" students. The 18 "analytic-tutorial" mean change on the A_1B_1 essays was negative, the students producing A_1B_1 posttest essays that were judged, on the average, 0.39 of a point lower in quality than the pretest essays. However, the posttest mean for the A_1B_1 essays was still larger by 0.15 of a point than that observed for all 180 E 306 students (see Table 4.1) and was only 0.02 of a point smaller than that recorded for all "analytic-conventional" students (see Table 4.6). In comparison to the 18 "analytic-conventional" students in the present comparison, however, these 18 "analytic-tutorial" students wrote A_1B_1 posttest essays that were judged 0.56 of a point better. On the A_2B_2 topics, the 18 "analytic-tutorial" students wrote essays which improved, on the average, by 0.72 of a point, a mean change identical to that of the 18 "analytic-conventional" students in the present comparison. That average gain is also identical to that recorded for all 180 E 306 students (see Table 4.1) but is 0.09 of a point higher than the average gain of all "analytic-conventional" students (see Table 4.6). While the mean gain is nearly identical to the mean gains recorded for these other three groups of students, the mean score for the A_2B_2 posttest essays is not. The means score of 4.89 is 0.30 of a point lower than the mean observed for all 180 students (see Table 4.1), and 0.48 of a point lower than that for all "analytic-conventional" students (see Table 4.1), and 0.48 of a point lower than that for all "analytic-conventional" students (see Table 4.6). However, the 4.89 average score is 0.34 of a point higher than the average gain of the 18 "analtyic-conventional" students in the present comparison.

To determine whether the two instructional methods--the "conventional" and the "tutorial"--had any effect on the posttest scores of the two groups taught the "analytic" curriculum in the present comparison, we again ran a series of analyses of covariance to compare the A_1B_1 and A_2B_2 posttest essay means of the two groups. As in previous such analyses, the pretest scores were entered into the analyses as covariates. As Table 5.24 and Table 5.25

Insert Table 5.24 about here.

Insert Table 5.27 about here.

---

indicate, significant effects were found only for the two indices measuring the use of final free modifiers in the A_2B_2 essays. Obviously, the "tutorial" method was more effective than the "conventional" in enhancing the use of final free modifiers. No other significant effects were observed, indicating that neither method was more effective than the other in causing the measured changes in the written products of the 36 students in the present comparison.

We also compared the performances of these 18 "analytic-conventional" students and 18 "analytic-tutorial" students on the objective measures of writing related behaviors. The pretest and posttest means of the two groups, together with T-tests of the mean changes for these objective measures, appear in Table 5.26 and Table 5.27. These tables indicate that the "analytic-conventional" students realized a significant average gain on the language mechanics subtest of the McGraw-Hill Writing Test, while the "analytic-tutorial" students achieved statistically significant mean gains on both the language mechanics subtest

---

Insert Table 5.26 about here.

Insert Table 5.27 about here.

---

and the total score. The average changes observed for neither group were significant for the paragraph comprehension subtest of the McGraw-Hill Reading Test or for the Daly-Miller Writing Apprehension Test.

Some noticeable differences between the two groups do, however, show up. First, the mean change of the 18 "analytic-tutorial" students on the language mechanics subtest was 0.79 of a point higher than that recorded for all 180 students (see Table 4.5), 1.07 points higher than the mean change for all "analytic-conventional" students in the present comparison. Second, whereas the mean change for the 18 "analytic-conventional" students on the sentence patterns subtest was a negative one, the mean change for the 18 "analytic-tutorial" students was positive. Although this change was not found to be statistically significant beyond the .05 level, it was slightly larger than that recorded for all 180 E 306 students (see Table 4.5), an average gain which was statistically significant beyond the .001 level of confidence. On the paragraph patterns subtest, negative mean change scores were recorded for both groups. A fourth difference between the two groups in the present comparison can be seen in the respective mean changes recorded for the total score on the McGraw-Hill Writing Test. On the total score, the 18 "analytic-tutorial" students registered a mean change 3.01 points larger than the average change of the 18 "analytic-conventional" students. This higher average gain was also 0.56 of a point higher than that recorded for all 180 E 306 students (see Table 4.5) and 1.18 points higher than that recorded for all "analytic-conventional" students (see Table 4.10).

On the paragraph comprehension subtest of the McGraw-Hill Reading Test, neither of the groups realized statistically significant changes, although both observed mean change scores were positive. The larger of the two mean change scores, the one for the 18 "analytic-conventional" students, was 0.14 of a point less than the average gain observed for all 180 students (see Table 4.5), a gain which was statistically significant beyond the .021 level of confidence. However, the mean change for the 18 "analytic-conventional" students was 0.54 of a point smaller than that recorded for all "analytic-conventional" students (see Table 4.10).

To determine whether any of the mean changes on the objective measures of writing-related behaviors for either of the two groups could be attributed to the effects of the instructional methods, we used analyses of covariance to compare the posttest means of the two groups, entering the pretest scores in the analyses as covariates. As Table 5.28 reveals, neither instructional method

Insert Table 5.28 about here.

can be judged more effective than the other in producing changes along the six dimensions of the objective measures of writing-related behaviors.

The results of Principal Comparison 4 do not allow a conclusion strongly favoring either method of instruction for the 36 students studying the "analytic" curriculum, even though we found the "tutorial" method more effective in promoting the use of final free modifiers in the A_2B_2 essays. In our judgment, unless an instructional method promotes higher posttest holistic scores on one or the other sets of essays, it cannot be singled out as more effective than another. In this comparison the incoming abilities of both groups of students presented some difficulties in the analyses. Notice, for example, the apparently large difference in incoming level of performance on the A_1B_1 essays. The pretest mean for the 18 "analytic-tutorial" students is 1.06 points higher than that for the 18 "analytic-conventional" students. When we compared the A_1B_1 posttest means of the "tutorial" students with those of all 180 students and all "analytic-conventional" students, we found them to be comparable. This may suggest that at the beginning of the semester the "analytic-tutorial" students had already reached their potential for writing on essay topics such as those represented by the A_1B_1 topics. If this be the case, we suspect that no instructional method would have produced larger scores until the students as a group had grown considerably, both intellectually and psychologically.

Principal Comparison 5: "Analytic-Tutorial" Students (N=18) Compared with "Synthetic-Tutorial" Students (N=18). This last of the 10 comparisons in the present study was designed to test for the effect of the curriculum when the teacher variable and the instructional variable--in this case, "tutorial" instruction without the computer-assisted component--were controlled. With the exception of two different groups of students, two different teachers, and a different instructional method, Principal Comparison 5 is identical to Principal Comparison 2. Both were designed to test for the same effect.

The pretest and posttest means for the A_1B_1 and the A_2B_2 essays for both groups are presented in Table 5.29 and Table 5.30. These tables also report T-tests of the mean change for both groups on both sets of essays.

Insert Table 5.29 about here.

Insert Table 5.30 about here.

As these tables indicate, the average word length of the A_1B_1 posttest essays for both groups was greater than the average length of the pretest essays. While the mean increase for the "analytic" students was not statistically significant, the mean increase of 84.89 words for the "synthetic" students was significant beyond the .001 level of confidence. This average increase in essay length was nearly 78 words greater than the average increase observed for the "analytic" students in the present comparison. But it was also considerably larger than those recorded for several of the other groups examined in the present study. This average gain in essay length was over 39 words longer than that recorded for all "synthetic" students (see Table 4.26), over 62 words greater than the average gain recorded for all "tutorial" students (see Table 4.19), and over 103 words greater than the mean change observed for all "analytic" students (see Table 4.25).

Large differences between the mean change scores on the A_2B_2 essays for the two groups in the present comparison was also observed.

As with the mean changes observed for the A_1B_1 essays, the average change for the "analytic" students was not statistically significant, while the mean change for the synthetic students was significant, again beyond the .001 level of confidence. The mean gain in essay length of 146.61 words for the "synthetic" students was nearly 98 words greater than the average change recorded for the "analytic" students in the present comparison. It was similarly greater than the mean changes recorded for three other groups not unrelated to the present comparison. The present "synthetic' students' mean gain in A_2B_2 essay length was about 59 words greater than that for all "synthetic" students (see Table 4.26), but it was nearly 93 words greater than that recorded for all "tutorial" students (see Table 4.19) and well over twice as large as the average gain for all "analytic" curriculum students (see Table 4.25). This large gain for the present "synthetic" students needs, however, to be viewed with reference to the average length of their A_2B_2 pretest essays. Although the average length of approximately 383 words for the present "synthetic" students' A_1B_1 pretest essays was only about 24 words smaller than that of their "analytic" counterparts, this average essay length was considerably smaller than those recorded for all "synthetic" students (see Table 4.26), all "tutorial" students (see Table 4.19), and all "analytic" students (see Table 4.25). Thus it would seem that at least part of the large gain for the present "synthetic" students can be explained with reference to their shorter pretests, as well as with reference to their longer posttest essays.

The "synthetic" and the "analytic" students in the present comparison also differed with respect to the gains measured by the syntactic indices. The changes recorded for both groups on the A_1B_1 essays in mean T-unit length are particularly noteworthy. Whereas the "analytic" students realized a mean gain of 0.35 words which was not statistically significant, the "synthetic" students managed a loss in mean T-unit length of 1.56 words. This large regression was significant beyond the .023 level of confidence, and it represented the only statistically significant loss in mean T-unit length for any of the several groups of students we examined in the present study. So great was the average loss for the "synthetic" students that their A_1B_1 posttest mean T-unit length, while 0.28 of a word longer than that of the present "analytic" students, was about a half a word shorter than the mean T-unit lengths recorded for all "analytic" students (see Table 4.25) and all "tutorial" students (see Table 4.19) and about three-fourths of a word shorter than the mean T-unit length of all "synthetic" students (see Table 4.26). Both of the groups regressed somewhat in average clause length on the A_1B_1 essays, but the posttest mean for the "synthetic" students was about a half of a word larger than that of the present "analytic" students.

Some interesting differences between the two groups also appear when the changes in mean T-unit and clause length for the A_2B_2 essays are compared. While the "analytic" group realized 0.28 of word greater gain than the "synthetic" students, the posttest mean for the "synthetic" group was over two words larger than that of the "analytic" group. It was also at least one word larger than the means recorded for all "analytic" students (see Table 4.25) and for all "tutorial" students (see Table 4.19) and about one half of a word larger than that observed for all "synthetic" students (see Table 4.26). The mean clause length for the A_2B_2 essays showed a very slight loss for the "analytic" students and a gain of 0.41 of a word for the "synthetic" students. This average gain for the "synthetic" students was about one word larger than that recorded for the "analytic" students in the present comparison. The "synthetic" posttest mean was also 0.36 of a word larger than that of all "synthetic" students, 0.65 of a word longer than that of all "tutorial" students, and 0.91 of a word longer than that of all "analytic" students.

Although the 18 "analytic" and the 18 "synthetic" students realized both comparable A_1B_1 posttest means and comparable A_1B_1 mean changes scores on the two indices designed to measure the use of final free modifiers, significant differences appeared between the two groups along these dimensions on the A_2B_2 essays. As Table 5.29 and Table 5.30 show, the "synthetic" students realized positive mean changes on both indices, while the "analytic" students regressed on both. In part a result of change in opposite directions for the two groups, the posttest means for the 18 "synthetic" students on both indices are substantially larger than those for the 18 "analytic" students. Converted to percentages, the ratios cited in the two tables indicate that the "synthetic" students used final free modifiers in 5.6% more of their T-units than the "analytic" students used in theirs. Again converted to percentages, these ratios indicate that these 18 "synthetic" students wrote final free modifiers in 1.4% more of their T-units than all "tutorial" students wrote in theirs (see Table 4.19), and in 3.2% more of their T-units than all "analytic" students wrote in their T-units (see Table 4.25). However, the mean percentage of A_2B_2 posttest

T-units for the 18 "synthetic" students was almost identical to that calculated for all "synthetic" students (see Table 4.26).

An examination of the mean holistic scores for both groups of students on both sets of essays also indicates some striking differences between the 18 "analytic-tutorial" and 18 "synthetic-tutorial" students. Table 5.29 and Table 5.30 show that on the A_1B_1 essays the "synthetic" students, on the average, increased their pretest scores by 0.89 of a point, or by 0.45 of a point more than did the "analytic" students. Although neither the average gain by the "analytic" students nor the average gain by the "synthetic" students reached statistical significance, the mean gain for the "synthetic" students exceeded the statistically significant gains of other groups examined in the present study. The 0.89 of a point gain, for example, exceeded by 0.41 the mean gain for all "analytic" students (see Table 4.25), by 0.65 that of all "tutorial" students (see Table 4.19), and by 0.35 the mean gain of all "synthetic" students (see Table 4.26). But while the mean gain in A_1B_1 essay quality for the "synthetic" students is somewhat impressive, their A_1B_1 posttest mean score is not nearly so. That score of 4.44 points is 0.23 of a point less than that recorded for all "synthetic" students (see Table 4.26), 0.34 less than the posttest mean of all "tutorial" students (see Table 4.19), 0.73 less than the mean of all "analytic" students' posttest mean, and 0.62 of a point less than the mean holistic score observed for the 18 "analytic" students in the present comparison.

Whereas the mean gain score on the A_1B_1 essays was larger for the "synthetic" students, the mean gain on the A_2B_2 holistic score was larger for the "analytic" students. In fact, the "analytic" students on the average improved by one full point on the A_2B_2 posttest, an average gain which is exactly twice as large as that observed for the "synthetic" students. This average improvement in essay quality is 0.17 of a point larger than it was for all "tutorial" students (see Table 4.19), 0.30 larger than the one for all "analytic" students (see Table 4.25), and 0.22 larger than the average gain observed for all "synthetic" students (see Table 4.26). But the differences between the mean changes on the A_2B_2 holistic score are not the only differences which were found along this dimension. The A_2B_2 posttest mean holistic score for the 18 "synthetic" students was 4.33; and for the 18 "synthetic" students, it was 5.44. The difference between the two posttest means is thus 1.11, indicating that at the end of the semester the "analytic" students were writing argumentative essays having a significantly higher quality than the essays of the "synthetic" students. Not only did these 18 "analytic" students write substantially better posttest argumentative essays than did the 18 "synthetic" students but so too did at least three other groups of E 306 students we examined. The 18 "synthetic" students wrote posttest argumentative essays which were inferior in quality to those produced by all "synthetic" students (see Table 4.26), by all "tutorial" students (see Table 4.19), and by all "analytic" students (see Table 4.25). These large differences were 0.58, 0.67, and 0.98, respectively.

To answer the question of whether either of the two curricula taught in a "tutorial" setting had a significantly greater effect on the posttest means of the 18 "analytic" and the 18 "synthetic" students, we compared the posttest means of the two groups. For this comparison, we conducted a series of analyses of covariance, using the pretest scores of each group as covariates in order to accommodate any pretest differences between the two groups. The results of these analyses are reported in Table 5.31 and Table 5.32.

Insert Table 5.31 about here.

Insert Table 5.32 about here.

Table 5.31 indicates that when pretest differences are entered into the analytic paradigm, neither the "synthetic-tutorial" course nor the "analytic-tutorial" course is more effective than the other in producing differences between the pretest and the posttest scores for the A_1B_1 essays. For the A_2B_2 posttest essays, however, the analyses of covariance calculated two statistically significant curricular effects and suggested a third. As an examination of the posttest means recorded in Table 5.32 reveals, the "synthetic" curriculum taught tutorially produced average essay lengths that were almost significantly larger than those produced by the "analytic" curriculum. The posttest differences in the use of final free modifiers were, however, the result of statistically significant curriculum effect. As would be expected, given the attention to sentence-level operations, the "synthetic" students used final free is somewhat smaller than that observed for all "tutorial" students (see Table 4.23) and somewhat larger than that recorded for all "synthetic" students (see Table 4.30). The mean change for the present "analytic-tutorial" students, on the other hand, is 0.82 of a point larger than that of all "tutorial" students and 0.88 of a point larger than the one observed for all "analytic" students (see Table 4.29). For the mean total scores for the several groups the differences do not appear to be very large.

A noteworthy difference between the two groups in the present comparison appeared for the respective change scores recorded for the paragraph comprehension subtest of the McGraw-Hill Reading Test. Whereas the "synthetic-tutorial" students realized a mean positive change of 1.29 points--which was 0.51 of a point larger than that recorded for all 180 students (see Table 4.5)--the "analytic-tutorial" students regressed by 0.43 points. This negative change compares with a positive change of 0.78 of a point for all 180 students (see Table 4.5), 0.58 of a point for all "tutorial" students (see Table 4.23), and 0.88 of a point for all "analytic" students (see Table 4.29).

Analyses of covariance allowed us to determine whether either of the two curricula taught tutorially could be said to better produce change along these dimensions than the other. As Table 5.35 reveals, significant effects were found for neither of the

---

Insert Table 5.35 about here.

---

two curricula, in spite of the large differences observed between certain of the objective test change scores recorded for the two groups.

5.3 The Five Derivative Comparisons: A Summary

Analyses of covariance conducted in connection with the five principal comparisons allowed us to focus more directly on the central questions posed by the present study. This question may be stated as follows: Is the regular E 306 course more or less effective than the alternative E 306 course? As was pointed out above, these two courses differed with respect to underlying rhetorical theory (an "analytic" or whole-to-part curriculum), instructional method ("conventional" classroom instruction in contrast to "tutorial," individualized instruction), and instructional media (textbooks in contrast to textbooks supplemented by interactive computer modules). In a direct comparison of these two E 306 options it was impossible to control or test for the instructional and curricular variables which distinguished the two E 306 options. However, a direct comparison--which controlled for the teacher variable and for the number of students on either side of the comparison--yielded some important results.

Analyses of the writing performances of the two groups of students on the narrative-descriptive essays that drew on personal experience indicate that the students in the "analytic" or regular E 306 course wrote posttest essays which were significantly better in quality than those written by the students in the "synthetic" or alternative E 306 course. These analyses also indicated that the "analytic" or regular E 306 course effected significantly larger number of T-units with non-restrictive modifiers in the final position than did the alternative, "synthetic" E 306 course. So too with the percentage of total words in final unbound sentence modifiers. A comparison of the posttest means for clause and T-unit length, however, revealed that neither course effected greater pretest-to-posttest changes along those dimensions of syntactic maturity than did the other.

As for the performances of the two groups of students on the argumentative essays, our direct comparison of the regular or "analytic" E 306 course and the alternative "synthetic" E 306 course indicates that neither course was more effective than the other in producing changes along the dimensions of mean clause length, mean percent of T-units with final free modifiers, mean percent of words in final modifiers, or mean quality score. Significant gains were observed along two dimensions. The regular "analytic" E 306 course produced longer essays than did the alternative E 306 course; and it produced posttest T-units which were, on the average, significantly longer than those produced by the alternative E 306 course.

No significant course effects were observed for the objective measures of writing-related skills. Both courses were equally effective in producing changes on the three subtests and the total score of the McGraw-Hill Writing Test, on the paragraph comprehension subtest of the McGraw-Hill Reading Test. Similarly, neither course effected any change in writing apprehension.

Our direct comparison of the two E 306 options could not control for the curriculum variable ("analytic" vs "synthetic" curriculum) or the instructional variables ("conventional" classroom instruction vs "tutorial" instruction and textbook instruction vs textbook instruction supplemented by computer-assisted instruction). In other comparisons, however, we did control for these variables, as well as the teacher variable and the number of students on either side of a given comparison. These other comparisons allow us to identify some of the most and some of the least salient features of the two E 306 options.

To determine whether either of the two curricula was more effective than the other, we compared the two curricula in two different ways. We compared the two curricula when both were offered in a "conventional" classroom setting and when both were offered in a "tutorial" setting. Along the eighteen dimensions we analyzed for a curriculum effect when the curriculum was taught in a "conventional" classroom setting, we found that the "analytic-conventional" courses were, with one exception, as effective as the "synthetic conventional" course. Students in the "synthetic-conventional" courses wrote posttest essays on the argumentative topics which were significantly longer than those of the students enrolled in the "analytic-conventional" courses. Our second comparison of the two curricula differed from the first in that both were offered in a "tutorial" setting rather than in a "conventional" classroom setting. Again we tested for the effects of the curricula on student performances along the 18 dimensions associated with the two types of essays and the objective measures of writing-related behaviors. Our analyses revealed that when taught in a "tutorial" setting, the "synthetic" curriculum caused students to write argumentative posttest essays that included a significantly larger mean percent of T-units with final nonrestrictive modifiers. The analyses also revealed that these same students in their posttest argumentative essays placed a significantly larger percentage of total words in these final free modifiers. Along the remaining 14 dimensions, neither curriculum taught in a "tutorial" setting could be judged more effective than the other.

The three statistically significant course effects attributed to the "synthetic" curriculum when the instructional variables were controlled, however, need to be placed in a context. While the "synthetic-conventional" students wrote significantly longer posttest argumentative essays than the "analytic-conventional" students did, there is little evidence to suggest corresponding or complementary significant posttest differences for argumentative essay quality. Similarly, although the "synthetic-tutorial" curriculum produced statistically significant effects associated with final nonrestrictive modifiers, the "synthetic-tutorial" courses did not prove to be more effective than the "analytic-tutorial" courses in enhancing the overall quality of the posttest argumentative essays. Without corresponding gains in writing quality, such increases in quantitative measures may be of little importance.

In addition to testing for the effects of the curricula when the other major variables were rigorously controlled, we also tested for the effects of instructional method while controlling for the other major variables. This comparison involved four "analytic" curriculum classes, two taught in a "conventional" classroom setting and two taught in a "tutorial" setting. Our analysis revealed that the "tutorial" method effected statistically significant posttest differences on the argumentative essays along the two dimensions focusing on the use of final nonrestrictive modifiers. Along the remaining 14 dimensions, no effects were found that could be attributed to the instructional method. Thus, along 14 of the sixteen dimensions both instructional methods appear equally effective in teaching the "analytic" curriculum.

The extent to which these more rigorously controlled comparisons shed light on our direct comparison of the two major E 306 options is open to question. Nevertheless, it is our opinion that these additional comparisons, together with the previously reported derivative comparisons, help to explain why the "analytic-conventional" E 306 option is in some ways more effective than the "synthetic-laboratory-tutorial" E 306 option. Our analyses of covariance did allow us to test systematically for the effects of curricula and instructional methods on the perfromances of E 306 students. We did not, however, systematically test for the effect of the computer-assisted instruction used as the "laboratory" component of the "synthetic-tutorial" E 306 option. We found that the "analytic-conventional" E 306 option produced five effects that were statistically significant, three on the posttest narrative-descriptive essays and two on the posttest argumentative essays. In all other respects the two principal E 306 options were found to be equally effective. When, however, the "computer-assisted" components of the "synthetic-tutorial" E 306 option was eliminated, the "synthetic" curriculum produced significantly longer posttest argumentative essays than did the "analytic" curriculum, and it promoted significantly greater use of final free or nonrestrictive modifiers in the posttest argumentative essays. Furthermore, in both of the comparisons of the two curricula when the institutional method was rigidly controlled, the "synthetic" curriculum produced essays as high in quality, on the average, as those produced by the "analytic" curriculum, and on both types of posttest essays. In our direct comparison of the two E 306 options, we found that the "analytic" option produced better essays in the posttest "narrative-descriptive" essays that drew on personal experience; and in our comparison of the two instructional methods--the "conventional" classroom and the "tutorial"--we found that both produced posttest essays on both topics that were equal in quality. Therefore, we must conclude that with reference to the direct comparison of the two major E 306 options, the "computer-assisted" component --not the "synthetic" curriculum and not the "tutorial" method affected adversely the quality of writing on the posttest narrative-descriptive essays. This component did not, however, have a similar adverse effect on the quality of the posttest argumentative essays.

CHAPTER 6

THE INFLUENCE OF WRITING APPREHENSION ON STUDENT PERFORMANCE IN E 306

In this chapter and the next chapter, we will examine factors which influence writing performance. The present chapter focuses upon how college freshmen feel about writing and how writing apprehension affects what they write. From classical antiquity onward teachers have observed that some individuals dislike writing and avoid situations where writing is required. Other individuals, in contrast, enjoy writing. They find writing a gratifying activity, even without the presence of external rewards.

6.1 Summary of Previous Research

Only recently have researchers examined how an individual's attitude toward writing might affect what he or she writes and how he or she reacts to situations associated with writing. Daly and Miller (1975a) call this phenomenon writing apprehension. They define writing apprehension as an individual's inclination to approach or avoid situations which are perceived to require writing accompanied by some amount of evaluation. Writing apprehension has been linked to several other factors, including those involving decisions concerning academic and career choices, performance on tests of verbal intelligence, and competency in writing. High writing apprehensives tend to choose academic and career specialties which they perceive to require little writing. Low writing apprehensives, on the other hand, tend to select occupations and majors which they believe to require considerable writing (Daly & Shamo, 1976, 1977; Miller & Daly, 1975b). The degree of writing apprehension has been shown to be associated with lower scores on standardized tests of writing aptitude and general verbal ability (Daly, 1978b; Daly & Miller, 1975b). More important to the present study, individuals who are highly apprehensive about writing produce essays which are evaluated lower in quality than those written by low apprehensives (Daly, 1977; Richmond & Dickson, 1980).

6.2 E 306 Writers and Writing Apprehension

Because of the results of these and other studies of writing apprehension, we decided to include the writing apprehension instrument as one of the measures in our study of E 306. As we reported in Chapters 4 and 5, none of the instructional and curricular components we analyzed for E 306 effected significant change in the degree to which E 306 writers were apprehensive about writing. However, we felt that the writing apprehension instrument measured an important component of students' receptiveness to writing instruction. We examined this assumption by testing four hypotheses: at the time of the pretest, high writing apprehensives would write essays that (1) were shorter,

(2) were less elaborated syntactically, and (3) would be evaluated by writing teachers lower in overall quality than essays written by low apprehensives; and (4) high writing apprehensives would score lower on standardized tests of writing ability, reading comprehension, and general verbal ability than low apprehensives.

Method. The 167 students who completed the writing apprehension instrument at the time of the pretest were divided into thirds on the basis of their scores on the instrument. We classified 55 students as high writing apprehensives, 51 as moderate apprehensives, and 55 as low apprehensives. We compared mean scores for the high and low apprehensives on the A_1B_1 essay topics, the expository topics drawing on personal experience, and the A_2B_2 topics, the persuasive topics, in an analysis of variance (ANOVA). We included total length, four syntactic variables (words per T-unit, words per clause, ratio of T-units with final nonrestrictive modifiers, and ratio of total words in final nonrestrictive modifiers --all used in analyses reported in Chapters 3, 4, and 5--and holistic rating of quality. We also compared mean scores on standardized tests of writing ability (McGraw-Hill Writing Test, English Composition Test), of reading ability (McGraw-Hill Reading Test, paragraph comprehension subtest; SATV-Reading Comprehension; and SATV-Vocabulary) and general verbal ability (SAT Verbal) in an analysis of variance. In addition, we attempted to determine which variables are predictors of writing apprehension by placing all dependent variables in a stepwise multiple regression analysis, a statistical procedure which estimates an independent variable in terms of a number of dependent variables.

Results. The results of the ANOVA of the A_1B_1 topics are reported in Table 6.1. As this table shows, low apprehensives significantly outperformed the high apprehensives on all essay measures at or beyond the .04 level of confidence.

Insert Table 6.1 about here.

Low apprehensives wrote papers which were, on the average, 83 words longer, more elaborate and varied in syntax, and higher in a subjective rating of overall quality.

The high and low apprehensives were not so clearly differentiated on the A_2B_2 topics. The results of this comparison are reported in Table 6.2. Table 6.2 shows that on the A_2B_2 essays the low apprehensives again wrote, on the average, papers significantly longer than high apprehensives and used significantly more non-restrictive modifiers. However, the differences between the high and low writing apprehensives in mean words per T-unit, mean words per clause, and mean holistic score were not significant.

Insert Table 6.2 about here.

Scores on standardized tests of writing aptitude, reading comprehension, and general verbal ability also differentiated the two groups, as Table 6.3 illustrates.

Insert Table 6.3 about here.

Differences between the mean scores of high and low writing apprehensives on the English Composition Test were significant at the

.001 level of confidence, but the scores on all of the subtests of the McGraw-Hill Writing Test (not reported here) and the total scores were not significant. Low apprehensives also performed significantly better on the Test of Standard Written English. Tests of reading comprehension showed a similar pattern, with low apprehensives outperforming high apprehensives at or beyond the .03 level of confidence on the McGraw-Hill Reading Test, the SAT Vocabulary subtest, and the SAT-Verbal test.

All 167 students were considered in the stepwise multiple regression analysis to determine which variables predict writing apprehension. These results are reported in Table 6.4.

Insert Table 6.4 about here.

The best predictor of writing apprehension was the number of words per clause on the A_1B_1 set of essays, explaining over 11 percent of the variance in scores on the Writing Apprehension Test. The total number of words in the A_1B_1 essays accounts for another 6.7 percent, and the ratio of words in final nonrestrictive modifiers in the A_2B_2 topics explains another 3 percent. The McGraw-Hill Reading Test, the only standardized test that explains as much as one percent of the variance in apprehension scores, accounts for three percent more. Altogether, the 18 variables entered into the regression equation account for about 28 percent of the variance in writing apprehension scores, leaving about 72 percent of the variance unexplained.

Discussion. Our analyses of the performances of high and low writing apprehensives support and extend the findings of others who have investigated writing apprehension. Writing apprehension again is associated with measures of general verbal abilities and writing performance. For the first time, writing apprehension has been linked to syntactic measures in written discourse.

Several findings are of particular interest. The fact that high writing apprehensives comparatively did much better on the A_2B_2 essay topics, the persuasive topics, suggests that these writers perform better on subjects not associated with the self. This interpretation is consistent with research that has associated writing apprehension inversely with general self-esteem (Daly, 1980). Topic variables thus seem important in assessing the writing abilities of high writing apprehensives.

The syntactic differences uncovered in this study for the first time point to a tendency of highly apprehensive writers to elaborate their statements less fully and to use the same syntactic constructions more frequently than low apprehensives. This may be attributable to generally lower verbal abilities as the standardized tests suggest, or it may in part be due to the tendency of these writers to "play it safe," using only words and structures they know how to spell and punctuate. Evidence exists that high writing apprehensives experienced negative teacher responses in their early writing attempts (Daly, 1978b; Harvley-Felder, 1978). Whatever the case, the multiple regression analysis shows syntactic characteristics much more closely related to writing apprehension than to overall writing quality (see Chapter 7).

6.3 Conclusion

Writing apprehension has again been shown to be an important dimension of writing performance, one which standardized tests do not measure to any appreciable degree. These and other results indicate that for college writing programs to be successful in teaching all students, they must reduce the anxiety about writing of highly apprehensive individuals. How this might be accomplished, however, has yet to be even speculatively discussed. Certainly, more research in the effects of instruction on writing apprehension is needed.

CHAPTER 7

PREDICTORS OF WRITING QUALITY AND COURSE GRADE IN E 306

In the preceding chapter we explored the relationship between writing quality and writing apprehension, two of the variables which we had measured for most of the 180 E 306 students used in the present study. In the present chapter, we explore most of the other variables analyzed in the preceding chapters to determine how much they predict of the variance in either the holistic scores or E 306 course grades.

7.1 Summary of Previous Research

In determining instructional emphases in the courses they teach, many writing teachers are guided by what they perceive as elements of writing assoicated with writing quality. As a result, some teachers heavily emphasize, say, syntactic features as in sentence combining, while others heavily emphasize, say, paragraph development. However, apart from a large body of opinion on the subject, there has been generated over the years very little empirical evidence to indicate which particular variables predict writing quality; and no studies, to our knowledge, have attempted to identify systematically those variables which predict overall performance in a freshman writing course.

Although relatively few studies have systematically examined the relationship between features of written discourse and raters' judgments of overall quality, some important work has been done. For example, Diederich, French, and Carlton (1961), in an analysis of teachers' comments on student papers, found that such matters as mechanical correctness, handwriting, organization, and tone seemed to influence the overall judgment of quality. Diederich, French, and Carlton also indicated that essay length affected raters' judgments of overall qulaity. Several years later, Nold and Freedman (1977), using more sophisticated statistical procedures, found essay length to be an important predictor of quality judgments. Using a stepwise multiple regression analysis, Nold and Freedman also found that, in their pre- and posttest sample of two types of discourse written by 22 Stanford University freshmen, the percentage of words in final nonrestrictive modifiers predicted 12% of the variance in the holistic scores. They also found that certain types of verbs and verb forms predicted another 11% of the variance. In a recent study of narrative and descriptive writing of University of North Dakota freshman, Faigley (1979c) reported that the percentage of T-units with final free modifiers predicted 16% of the variance in the holistic scores assigned to the pretest and posttest essay of his sample population. Essay length was found to be the second most significant predictor, but it predicted only about 3% of the variance in holistic scores. In the Faigley study, a total of 22% of the variance in the summed holistic scores was predicted by the sums of the six variables entered into the regression equation.

7.2 Predictors of Writing Quality in E 306

To determine how much of the variance in holistic scores of the 720 E 306 essays (see Chapter 2) used in the present study could be predicted by the five variables computed for each essay set, we ran four stepwise multiple regression analyses. The appropriate holistic scores were used as the dependent variable in each analysis; and the independent variables in each analysis were the four syntactic variables--mean T-unit length, mean clause length, ratio of T-units with final nonrestrictive modifiers to total T-units, and the ratio of words used in final free modifiers to total words--and the variable for average essay length. A stepwise multiple regression enters into the regression equation the independent variables in the order of their predictive power. It then continues entering independent variables into the equation until the researcher decides from the entering F-ratio that the added variables are not significant predictors of the dependent variable.

The first of the four regression analyses we conducted for holistic scores focused on the relationship between the quality scores assigned to be A_1B_1 posttest essays and the five independent variables listed above. The results of this analysis appear in Table 7.1. As Table 7.1 indicates, approximately 5% of the

---

Insert Table 7.1 about here.

---

variance in the holistic scores assigned to the A_1B_1 posttest essays is predicted by average essay length. Another 3% is predicted by the ratio to total words of words in final free modifiers. Collectively, all five variables predict less than 8 1/2% of the variance in holistic scores, leaving over 91% of the variance unexplained.

Table 7.2 presents a similar analysis, this one for the

Insert Table 7.2 about here.

---

holistic scores assigned to the A_2B_2 posttest essays. Of the five independent variables entered into the regression equation, only one--mean essay length--was found to predict more than 1% of the variance in the holistic scores. That variable predicted a little over 6.5% of the variance. The other two variables listed in Table 7.2 together predict less than 1 1/2% of the variance in holistic scores. All five variables entered into the equation predict collectively only slightly more than 8% of the variance, leaving over 91% of the variance unaccounted for.

In the third stepwise multiple regression, we used the summed holistic scores of the A_1B_1 and the A_2B_2 posttest essays as the dependent variable and the eight syntactic factors and the two essay-length factors as the independent variables. As Table 7.3 illustrates, the best predictor of the posttest

Insert Table 7.3 about here.

holistic scores is the average number of words in the A_2B_2 posttest essays. That variable predicts nearly 8% of the variance. The next best predictor--ratio of words in final nonrestrictive modifiers in the A_1B_1 posttest essays--predicts an additional 4.5% of the variance. The remaining eight factors each predict less than 1% of the variance, and all ten independent variables collectively predict approximately 14.3% of the variance in the summed holistic scores of the A_1B_1 and the A_2B_2 posttest essays.

To conduct the fourth stepwise multiple regression analysis, we summed the holistic scores for the pre- and posttest essays written on both the A_1B_1 and the A_2B_2 topics. The resultant summed holistic score--which was the sum of the four holistic scores for each case--was used as the dependent variable in the regression equation. For the independent variables, we summed the four scores for the mean number of words, mean T-unit and clause length, and the two ratios associated with final free modifiers. The results of this analysis are reported in Table 7.4. As Table 7.4 shows, essay length was again found to be the

Insert Table 7.4 about here.

best predictor of writing quality. This factor predicted a little more than 11% of the variance in the summed holistic scores. The second best predictor--ratio of words in final nonrestrictive modifiers--predicted only 2.3% of the variance; and words per T-unit--the third best predictor--accounted for less than 1 1/2% of the variance. Unlike the other two predictors, however, mean T-unit length is negatively associated with writing quality, as the correlation coefficient (r= -.0060) for the two variables indicates. Collectively, all five independent variables explained only 15.02% of the variance in the summed holistic scores of the two pretest and two posttest essay sets.

7.3 Predictors of Course Grades in E 306

Of great interest to writing teachers and administrators alike are measures that can predict success in a writing course. Teachers and administrators are frequently told that standardized tests such as SAT correlate highly with course grades in writing classes, but seldom are they told that causality cannot be inferred from correlation coefficients. That is to say, writing teachers and administrators often assume that a causal relationship exists between, say, SAT Verbal scores and semester grades in a writing course. The extent to which such causal relationships obtained for the 180 E 306 students used in the present study is the subject of the present section.

To identify significant predictors of semester grades in E 306, we ran a stepwise multiple regression analysis, using final course grades in E 306 as the dependent variable. For this analysis, we used as independent variables several objective measures of writing-related skills. The objective measures of writing-related skills used as independent variables were the following: SAT Verbal, SAT Math, English Composition Text (ECT), and the three subtests of the McGraw-Hill Writing Test--language mechanics, sentence patterns, and paragraph patterns. (For an exposition of the three McGraw-Hill subtests, see Chapter 4 above.) We also used as independent variables the A_1B_1 and the A_2B_2 pretest holistic scores, plus the one essay-length index and the four syntactic indices computed for both the A_1B_1 and the A_2B_2 pretest essays. Although some of the data used to compute the independent variables were collected several months prior to the students' matriculation in E 306 and some were collected at the very beginning of their E 306 study, none of the data was collected after the students received instruction in a college writing course.

Table 7.5 reports the results of this analysis for the four most powerful predictors of E 306 course grade. As Table 7.5 shows, the pretest holistic score on the A_1B_1 essays is the best predictor of E 306 course grades, accounting for 5.8% of the variance among the final grades reported on the grade sheets at the end of the Fall Semester of 1978. The second best predictor--SAT Verbal--explained another 3.3% of the variance; and the ratio of T-units with final nonrestrictive modifiers in the

---

Insert Table 7.5 about here.

A_2B_2 pretest essay accounted for another 1.3%. The fourth most powerful predictor--the ratio of words in final free modifiers in the A_1B_1 pretest essays--explained slightly less than 1% of the variance. It should be noted, however, that a negative correlation coefficient ($r = -.0286$) was computed for the relationship between course grade and ratio of words in final nonrestrictive modifiers. The remaining twelve independent variables combined to explain only an additional 4.9% of the variance in E 306 semester grades. Thus altogether the 17 independent variables entered stepwise into the regression equation explained only 16.24% of the variance in course grades, leaving 83.76% unexplained.

7.4 Discussion of the Results

The results of the analyses presented in this chapter seem to contradict in some ways the findings reported in previous studies. By comparison to the percentage of variability in holistic scores reported by Nold and Freedman (1977) and by Faigley (1979c), the present percentages seem quite low. Nold and Freedman reported that a little over 31% of the variance in holistic scores was explained by two independent variables. One of these resembled the essay-length variable used in the present study. The other was a simple conversion to percentage of the present study's ratio of words in final free modifiers. In the Nold and Freedman study, the latter variable was found to predict 11.6% of the variance in holistic scores. In the present study, the most that that variable predicted was about 4.5% of the variance in the holistic scores assigned to all posttest essays. The Faigley study, like the Nold and Freedman study, found one index of final nonrestrictive modifiers to be a powerful predictor of holistic score. Faigley's index--the percentage of T-units with final free modifiers--is simply a conversion of the ratio used in the present study. While Faigley reported that 16% of the variance in holistic scores was explained by this index, the largest percentage of the variance in holistic scores found in the present study for that index was .65% on the A_2B_2 essays. As a predictor of E 306 course grade, this variable explained only about 1.3% of the variance. The other index of free-modifier use predicted even less of the variance--about 0.9%--and that variable correlated negatively with course grade.

Attempting to reconcile these different percentages explained by these variables is no easy matter. While we cannot identify precisely the reasons for these discrepancies, we can at least speculate about some of them. In the Faigley study, the regression analysis used the summed holistic scores awarded the pretest and the posttest essays. Of the students in the Faigley study, approximately 51% underwent intensive instruction in the principles and practice of Christensen's generative rhetoric. Thus fully 25% of the essays making up the sample on which the regression analysis was run were written by writers who had been taught systematically to add free modifiers to the base clauses of the setnences they wrote. In contrast, the "synthetic" students in the present study, students also taught a Christensen-based curriculum, numbered only 54, accounting for about 30% of the students used in the present study. Thus in the present study, Christensen-influenced posttest essays accounted for only about 15% of the total essays. This relatively smaller contribution in the present study of the "synthetic" essays to the total essays probably influenced the outcome of the regression analysis.

Another probable reason for the discrepancies in the predictive power of the two indices associated with final nonrestrictive modifiers can be explained with reference to holistic change scores of the "analytic" students in the present study and the holistic change scores of the "control" students in the Faigley study. In the Faigley study, the "control" students were taught a curriculum similar, although certainly not identical, to the "analytic" curriculum in the present study. In the Faigley study, the "control" students realized an average holistic change of 0.18 of a point on a 1-to-6 scale, gains which were not statistically significant. In constrast, the "analytic" students in the present sutdy realized statistically significant positive change in holistic score on both kinds of writing investigated in the present study. This difference between the two groups of "analytic/control" students indicates that in the Faigley regression analysis, the essays which were rated the highest were those of students who had been taught to use final nonrestrictive modifiers. In the present study, on the other hand, both the "synthetic" and the "analytic" students realized significant positive changes in holistic scores, thus insuring that the regression analyses would have been influenced both by students taught to use final nonrestrictive modifiers and by students who had not been taught to use final nonrestrctive modifiers, but had nevertheless improved the overall quality of their writing.

Another possible way to account for the differences is with reference to the training procedures used in holistic evaluations. In each of the three studies reviewed here, the raters were trained with essays believed representative of the essays to be rated in the respective holistic sessions. That is to say, the criteria or cues which the respective groups of raters used to judge the essays were always made relative to the essays of the population being judged. Unless we are prepared to believe--without the benefit of any supportive evidence to that effect--that the raters in each of the three studies were utilizing the same set of judgmental cues and weighting those cues similarly in their judgments of writing quality, we have to assume that the judgments of writing quality were made differently in each of the three studies. If this assumption is a valid one, then the predictive power of a given variable should vary from one study to another. And it appears to do so.

Even if the results of the three studies with reference to predictors of writing quality are not strictly comparable, the present study does in some sense confirm the validity of using essay-length indices and indices of free-modifier use in such multiple regression analyses. But perhaps more importantly, it squarely points to the need for developing rating procedures which are reliable across different studies and different populations. Together, the three studies we have reviewed here indicate the need to use a variety of new variables to determine the nature of writing quality more precisely. Our own study in Chapter 3 of the differences between high and low quality E 306 essays suggests some such indices to us. For example, some measure of cognitive egocentricism--such as the use of first- and second-person pronouns--seems to be a promising index for certain types of writing. The frequencies of error types seems to be another. Although extremely time-consuming, analyses of cohesion would probably yield variables capable of predicting a high percentage of the variance in holistic scores. The category of lexical cohesion seems particularly promising in this regard because it provides some measure of the extent to which writers are able to extend the semantic domains of concepts and ideas introduced in a piece of writing.

Our regression analysis to predict, on the basis of precourse data, the final course grades in E 306 also points to some interesting findings. Most importantly, our analysis shows that for the sample population used in the present study, several of the "standardized" test measures are totally ineffective in predicting success in a writing course. Of the five such measures we examined--SAT Verbal, SAT Math, ECT, and the McGraw-Hill Writing Test--only the SAT Verbal proved powerful enough to predict a significant percentage of the variability in course grades. And that variable predicted only a little more than 3% of that variability. This finding suggests that a good deal goes into the computation of semester grades which is not measured by the SAT Verbal. Similarly, our analysis indicates that with respect to writing quality either (1) composition teachers assign grades to essays differently than holistic raters do, or (2) writing quality does not have a significant impact on course grades. We suspect that the former is the proper explanation. This suspicion is grounded on the differences discussed previously in some detail in Chapter 2. It seems to us that teachers assign essay grades and course grades based to some extent on the relative performances of the group of students in the particular class. "Class" populations are, in contrast, irrelevant in holistic evaluations. It also seems to us that composition teachers often grade student writing according to different sets of criteria established for different writing assignments, perhaps emphasizing or weighting, say, mechanics more heavily on one assignment than another. Another major difference is that different criteria obtain for different types of writing. In composition classes themselves, students seldom write two comparable essays during the semester. Each assignment probably differs from the preceding and the following one with respect to mode, purpose, audience, or subject matter. Perhaps all assignments differ along four dimensions or several combinations of dimensions. Differing essays are and must be graded differentially.

At any rate, the present chapter suggests the need to identify and define variables which can predict performance in freshman writing classes better than the variables used in the present study. Only when better predictors of course achievement are found can directors of freshman composition programs have confidence in their placement procedures.

CHAPTER 8

THE PRESENT STUDY IN THE LARGER CONTEXT

The present study was undertaken for the specific purpose of comparing the relative effectiveness of two distinct freshman composition courses at the University of Texas. In one sense, then, both the design of the research and the results of that research are "local" in nature. However, the present study shares a larger context as well, for it addresses some major issues in writing program evaluation and in curricular and instructional design, issues which concern both the University of Texas in particular and other universities and colleges in general. The present chapter focuses on the local and non-local contexts of the research reported in the previous chapters.

8.1 The Larger Context of the Present Study

To place the present study in its larger context is to offer immediately a number of caveats, not the least of which concerns the state of writing program evaluation itself. Although educational researchers (e.g., Braddock, Lloyd-Jones, and Schoer, 1963; Gage, 1973; Mellon, 1975; and Larson, 1979) generally emphasize the importance of writing program evaluation, Larson (1979) contends that the evaluation of writing, of the teaching of writing, and of writing programs is, for the most part, "an art not yet born." Indeed, at the present time, the profession lacks not only common procedures for evaluating writing programs and courses but an adequate theoretical framework for developing such procedures as well (see Kinneavy, Daly, Faigley, and Witte, 1980). Difficulties in defining the characteristics of effective written discourse and in identifying the salient features of composition courses have kept knowledge of how to evaluate writing programs and courses at a rudimentary and impressionistic level. No one has yet articulated the shape such evaluations should take or the steps required to implement them. With no generally accepted and few widely applied or applicable procedures for evaluating writing programs, the larger context of the present study must remain somewhat elusive.

Within this elusive larger context, previous evaluation research has called attention to two related questions: Does college writing instruction positively affect the development of writing abilities? and Is one type of composition instruction more effective than another in teaching composition skills?

The University of Northern Iowa Study. Most studies have addressed both questions simultaneously, but one study focused exclusively on the first question. That study, conducted at the University of Northern Iowa (Jewell et al., 1969), paired 2,080 freshmen enrolled at five different universities in the fall of 1964. The paired students were matched on the basis of sex, writing ability, and ACT scores, with one member of each pair taking courses in writing and in other academic subjects and one member taking courses in other subjects only.

Jewell reports that three tests--English Expression (COOP), College Entrance Examination Board English Composition Test (CEEB), and a theme--were administered to the students in the study at four different times--once at the beginning of the college career and once at the end of the first, second, and fourth semesters. The analyses of the data indicated that at the end of the first two semesters the students undergoing instructional treatments in composition outperformed their paired counterparts. However, by the end of the fourth semester, students who had received no formal instruction in composition performed as well as those students who had received such instruction as freshmen. Jewell thus concluded that freshman composition teaches writing skills and behaviors that develop naturally, but at a later date. In one sense, the Jewell study would suggest that the composition instruction at the freshman level is of limited value. However, the assumption underlying the present study is that if freshman composition can accelerate writing development, students who have undergone composition instruction will benefit substantially during their early semesters in college from having received that instruction.

Other studies have focused principally on the second question: Is a particular type of freshman composition instruction more efficacious than another? Like the present study, those studies have assumed that writing instruction does benefit students. Many such studies have been conducted in recent years, and some are more important than others. Here we call attention to three we consider the most significant, even though all have certain design problems which render their finding inconclusive.

The University of California at San Diego Study. Among those studies is one recently completed at the University of California at San Diego (Wesling et al., 1978). The subjects used in that study were 175 freshmen unevenly distributed within four colleges within the university. Each of the four groups--which ranged in size from 35 to 50 students--produced three essays which were submitted to holistic evaluation by three independent raters. Each of the four groups had received instruction in composition, but apparently that instruction differed in important ways for each group. All groups are reported to have significantly increased their average holistic scores from the first to second themes and from the first to the third themes. Certain of the students also took the English Achievement Test (EAT) at two different times, once prior to admission and once after completing a course in college writing. The pretest-posttest use of the English Achievement Test was adopted in order to determine whether writing instruction in the four colleges "led to improvement in the editing component of writing" (p. 21). The gain scores on the English Achievement Test for none of the four college groups were statistically significant. Although the San Diego study is shot through with problems in research design (e.g., a lack of control over discourse variables such as mode and purpose in the three writing samples, the administration of the English Achievement Test under different conditions, and a lack of control over the subject matter variable in the essays which the students in the different colleges produced), the study does seem to suggest that writing instruction does have a positive impact on the writing performances of college freshmen. Two of the four college programs evaluated depended heavily on readings, one on readings from the humanities and one on readings from the social sciences, while the other two colleges appeared to offer composition instruction more focused on the actual essays the students produced. Yet all four groups of students apparently realized statistically significant improvement in their abilities to write essays of the kinds they were asked to write.

The Miami University Study. Another important study is the widely publicized and recently completed one at Miami University (Daiker, Kerek, and Morenberg, 1978; Morenberg, Daiker, and Kerek, 1978). This study, which is less complex in nature than either the Northern Iowa or the San Diego study, is an experimental study that attempts to determine the relative effectiveness of a "sentence-combining curriculum" in comparison to a composition course "traditional at Miami and many other universities." In the Miami University study, two groups of beginning college freshmen were formed. One was called a "control" group and one was called an "experimental" group. The former group was exposed to instruction based on McCrimmon's Writing with a Purpose supplemented by readings from an anthology of essays. The second group, the "experimental" group, was given instruction based on the sentence-combining exercises contained in Strong's Sentence Combining: A Composing Book. Both groups of students thus underwent an instructional treatment of some kind, thereby making the investigators' distinction between a "control" group and an "experimental" group a questionable distinction at best.

Although the reports of the Miami University sentence-combining experiment suggest the contrary, the instructional treatments offered the two groups differed along several instructional and curricular dimensions other than the sentence-combining exercises the "experimental" group completed. As Kinneavy (1979) and Mellon (1979) have pointed out, the sentence-combining exercises themselves were cast as whole-discourse problems, thereby teaching the "experimental" students rhetorical principles that may or may not have been taught the "control" group. Neither is there any indication that the students in the two groups were taught to write discourse similar with respect to aim or purpose and mode (on the distinction between aims and modes in written discourse, see Kinneavy, 1971; Kinneavy, Campbell, and Cope, 1976a and 1976b; and D'Angelo, 1980). Indeed, even the most cursory review of Strong's Sentence Combining, the text used in the "experimental" classes, reveals that narration and description are the two modes of discourse most frequently represented in the sentence-combining exercises, and that the purpose or aim of most of the exercises is to inform. In contrast, a "traditional" composition course would teach other modes as well as narration and description and would probably teach students to write effective persuasive, argumentative, and expressive discourse as well as informative discourse. Indeed, if the "control" classes in the Miami University study were taught the McCrimmon text in its entirety, they should have been taught to write in several different modes and for a variety of purposes.

The reports of the Miami sentence-combining experiment are silent on a number of other matters also. The writing topics used probably to elicit the pre- and posttest writing samples by which distinctions between the two groups were made called on the students to produce narrative and descriptive writing that drew heavily on personal experience, exactly the kinds of writing the sentence-combining students were taught to write throughout the semester. But was it the type of writing the "control" students undergoing "traditional" writing instruction were taught to write? We cannot tell from the published reports of that study. Neither do the Miami investigators indicate whether the students in the two groups wrote an equivalent, or even comparable, number of words during the semester, although the authors do report that both groups wrote the same number of essays at the same points during the semester. Yet the "experimental" or sentence-combining group combined sentences in preparation for or in connection with virtually every class meeting. They thus wrote on a much more regular basis than did the students in the "control" or "traditional" writing course.

The larger gains for the "experimental" students along the dimensions of holistic score and syntactic fluency also raise a series of questions. Were these larger gains in syntactic complexity and writing quality the direct result of sentence-combining practice? Were they the result of class discussions of students' solutions to sentence-combining problems? Were they the result of the inductive teaching of rhetorical principles? Were they the result of the great amount of writing that the sentence-combining students did? Were they the result of the classes' analyses of students' solutions to sentence-combining problems? Or were they the result of any possible combination of these uncontrolled variables in the experiment? In point of fact, it is impossible to tell from the accounts of this research exactly why the students taught sentence combining wrote better essays that were more syntactically complex than their counterparts in the "control" classes.

The University of North Dakota Study. Comparable to the Miami sentence-combining study is a study recently completed at the University of North Dakota (Faigley, 1979b, 1979c). It is comparable to the Miami study in at least two respects: first, it tested the effectiveness of a sentence-based approach to teaching composition against a "traditional" approach; second, it utilized a research design almost identical to the one used in the Miami study. However, it differs from the Miami study in important ways as well. First, it attempted to answer more questions than did the Miami study. Second, it controlled more rigorously some of the important curricular and instructional variables. Third, students in the "experimental" classes in the North Dakota study were taught generative rhetoric (see Christensen, 1967, for the major theoretical statement).

In contrast to sentence-combining--which is only an instructional method--"generative rhetoric" conflates instructional and curricular concerns. From an instructional point of view, generative rhetoric, like sentence combining, teaches students to write sentences which are syntactically more complex than the ones they customarily write. Both instructional methods encourage students to embed kernel sentences in independent clauses, to add subordinate elements to those clauses. Such embeddings allow writers to deliver more content in fewer words, reducing some independent clauses to phrases and others to subordinate clause status. The result of such training and practice is usually the production of sentences significantly longer than those the students were writing before they underwent the instructional treatment, whether sentence combining or generative rhetoric. While the addition of content to independent clauses is one of the principal goals of both sentence combining and generative rhetoric, there are other important differences between the two. Whereas sentence-combining exercises supply all of the content the students include in the derived sentences they write, instructional materials based on generative rhetoric frequently supply the student only the base clause for an extended sentence. The student subsequently adds his or her own content to the base clause.

Sentence combining and generative rhetoric also differ with respect to curricular concerns, what might be referred to as the "content" of a composition course. Sentence combining has no content per se, whereas generative rhetoric posits as content certain semantic principles which operate within discourse units, whether sentences or extended units of discourse such as paragraphs or essays (on this matter, see Chapters 2 and 3 above). A generative rhetoric curriculum teaches students to employ in the discourse they write specific semantic relationships or structures. Sentence combining, on the other hand, makes no assumptions about the nature of derived texts, whether sentences or larger units. Sentence combining refers simply to the act of combining kernel sentences to form larger derived sentences. It espouses no particular rhetorical theory, although exercises may be designed to teach certain rhetorical principles (on this matter, see Harris and Witte, in press). In contrast, the only theory that underlies sentence combining is a pedagogical one.

The Miami sentence-combining study speaks of sentence combining as though it were a curriculum, i.e., as though it offered a content as well as an instructional method. As a result, that study failed to control adequately either the major instructional or the major curricular variables. The North Dakota study of generative rhetoric, on the other hand, recognized the presence of both curricular and instructional variables in generative rhetoric; but it contained certain inherent limitations. The most important of these was the fact that the freshman English program at North Dakota was divided according to the traditional four modes. Narration and description were treated in the first semester and exposition and argumentation during the second semester. The study ended after the first semester; thus students were taught only narration and description. Narrative writing samples were collected. A follow-up study examining argumentative writing would have been desirable; however, the results of such a study could not be conclusive because the students were exposed to different instruction during the second semester. The North Dakota study failed to answer the criticisms of Johnson (1969), Tibbetts (1970), and Winterowd (1975) that generative rhetoric is of little use for teaching argumentation and exposition. Neither the Miami University nor the North Dakota study tested the effect of sentence-skills instruction on argumentative writing.

The North Dakota study examined the influence of generative rhetoric on specific syntactic structures as well as on the gross indices of syntactic complexity used in the Miami study and on writing quality. As in the Miami study, differences between the two groups' performances on a single writing sample were examined in detail. As the Miami study found statistically significant differences favoring the "experimental" classes at the end of the semester, so too did the North Dakota study. In addition to the "experimental" classes' larger gains in syntactic complexity and writing quality, students taught generative rhetoric wrote a higher percentage of their T-units with final nonrestrictive modifiers than did the "control" classes.

Although the San Diego and the Northern Iowa study both suggest the efficacy of "traditional" writing instruction at the college freshman level, they do so without the benefit of an adequate research design and without the visibility of a "gimmick." Using somewhat stronger research designs which controlled for a larger number of variables and offering a "gimmick" for teaching clearly focused writing courses, both the Miami study and the North Dakota study report findings and analyses which strongly suggest that "traditional" freshman writing courses (and traditional always seems to refer to courses not based on a "sentence-skills" approach) are inferior to courses based on sentence combining or generative rhetoric. That "traditional" courses in freshman composition should be rather poorly investigated--as in the San Diego study--or that they should be shown to be inferior to other types of freshman composition courses--as in the Miami University study and the North Dakota study--provides the larger context for the present study. In contrast to the Miami study and the North Dakota study, the present study compared two well-designed (and "well-designed" is an important construct generally ignored in previous research on "traditional" writing courses) and clearly focused freshman writing--one a "traditional" freshman writing course and the other a generative rhetorical course. The present study is thus distinguished from earlier published comparisons of "traditional" freshman writing courses with other courses because it controlled systematically for several curricular and instructional variables that were not controlled in previous research studies. The present study is thus far more notable for its research design and its control of major variables than any study published to date. For these reasons, the present study not only adds significantly to our knowledge of composition curricula and instructional methodologies, but it also serves as a model which can be used to guide future composition research, whether that research focuses on testing "experimental" approaches to the teaching of freshman writing or whether it focuses on comparative evaluations of extant approaches.

8.2 Summary of the Major Findings of the Present Study

The more rigorous control over major instructional and curricular variables in the present study allowed us to address with some confidence our major question. That question focused on the relative effectiveness of two distinctly different freshman writing courses. To investigate that question adequately, we could not, as several previous studies of college composition courses have done, assume that the two courses differed along one dimension only. As we pointed out in Chapter 2 above, the two courses in the present comparison differed with respect to underlying rhetorical theory and method of instruction. As a result, several curricular and instructional variables had to be controlled in order to determine whether either of the two courses was more effective than the other. Like the Miami University study, then, the present study compared two entirely different freshman writing courses. Unlike the Miami University study, however, the present study speaks of neither "control" groups nor "experimental" groups, treating both as though they were undergoing some type of "experimental" treatment, as indeed they were.

The present study also investigated the performances of the student groups more systematically than have previous evaluation studies. In the present study, for example, students in both courses did not write just one piece of discourse on which their performances were compared on the basis of two samples of extended written discourse mode and aim. As we noted in Chapter 2, one of the discourse samples was largely narrative and descriptive with respect to mode, drew heavily on the students' personal experiences, and was designed to be informative. The other discourse sample was argumentative in nature, comparable to what Kinneavy calls "scientific discourse," in which a writer attempts to prove a thesis. By eliciting two distinctively different types of written discourse from the students in the present study, we believed that we would be able to compare student writing on the basis of written products which are usually taught and written in "traditional" writing courses and which were specifically called for in the course syllabus for the Freshman English Program at the University of Texas.

In addition to comparing student performances on the basis of two writing samples representing different types of extended discourse, we also compared student performances on objective measures of writing-related skills or behaviors. Although these skills and behaviors were considered but adjuncts to actual writing practice, we believed that they would provide additional indicators of the impact of writing instruction. Specifically, we sought to measure the effects of writing instruction in five areas other than actual writing performance. First, we sought to determine the effects of writing instruction on the recognition of mechanical errors in punctuation, capitalization, and grammar. Second, we sought to measure the effects of writing instruction on the student's ability to recognize sentence patterns of various types. Third, we sought to measure how well a student could organize a given content in to a multiple-sentence paragraph and how well the student could recognize the appropriate and inappropriate use of transitions. Fourth, we sought to determine whether writing instruction had any effect on the student's ability to read and comprehend prose texts comparable to those he or she would encounter in textbooks for other college courses. Fifth, we tried to determine whether instruction in writing significantly increased or decreased the amount of anxiety or apprehension students experienced when confronted with writing tasks. Together with the holistic evaluation scores and the syntactic analyses of the writing samples, these five measures of writing-related skills and abilities constitute the multidimensional nature of the present study. Thus unlike most previous research in freshman writing, the present study compares the effectiveness of freshman writing courses not along one or two dimensions but along several dimensions of writing performance, writing-related skills, and the attitude toward writing.

It is neither possible nor desirable to summarize here all of the results reported in the previous chapters, but it is necessary to indicate those findings that we believe to be the most significant, significant both in terms of the larger context of composition research and in terms of the needs of the Freshman English Program at the University of Texas.

Among the most important results of the present study are the following ones. First, the present study shows that a *"synthetic," sentence-based course in freshman writing can be used to teach argumentative discourse effectively*--something no previous research in composition has indicated it could do. The "synthetic" curriculum in the present study was able to teach its students to write argumentative essays while requiring those students to write substantially fewer whole essays during the course of the semester than did the "analytic" curriculum in teaching students to write argumentative essays may be attributable to the greater development of certain invention skills among the students enrolled in that curriculum.

Second, the present study shows that *both the "analytic" and the "synthetic" curricula can be used effectively in both conventional "classroom" and "tutorial" settings*, another matter not addressed in previous research in freshman composition. Up to now, studies have assumed that either instructional or curricular variables were being investigated, but not both.

Third, the present study strongly suggests that *well-designed and clearly focused freshman writing courses can be equally effective in producing positive change in writing performance and in objectively measured skills and abilities.* This seems to be true regardless of whether the courses are based on a "synthetic" or an "analytic" curriculum and whether they are taught in a conventional "classroom" setting on in a "tutorial" setting. This finding is particularly important because it calls into question the findings of several recently published studies reporting the greater effectiveness of "sentence-based" approaches as compared to "traditional" approaches to teaching freshman writing.

Fourth, *neither the two curricula nor the instructional methods examined in the present study either reduced or increased significantly the amount of apprehension students feel when they approach writing tasks*. Yet, as reported in Chapter 6, writing apprehension does appear to influence the way students write and the way they perform on objective measures of writing-related skills. This important dimension in the teaching of freshman writing has never before been systematically investigated with reference to so many curricular and instructional variables.

Fifth, the results of paired T-tests of gain scores indicate that *students in virtually all the classes studied realized statistically significant improvement in not only overall writing quality as measured by holistic scores on two different measures of writing performance, but they also realized significant gains in their command of the conventions of Standard Written English as measured by the McGraw-Hill Writing Test*. Rarely have writing-related skills been considered in evaluations of writing curricula and instructional methods.

A sixth major finding of the present study is that virtually *all of the "analytic-conventional" classes realized statis-*

tically significant improvement in reading comprehension, a finding which no other major study of freshman composition courses has been able to show. Most recent attempts to link improvement in reading comprehension with writing instruction have focused on the "sentence-skills" approach (e.g., Morenberg, Daiker, & Kerek, 1978). Research in comprehension, on the other hand, has largely abandoned sentence-level approaches and now considers the influence of "top-level" structures, structures that organize an entire piece of discourse (e.g., Kintch & van Dijk, 1978; Meyer, 1975; Meyer & Freedle, 1979). Recent comprehension research supports the assumption that the analytic curriculum, with its emphasis on the structure of whole pieces of discourse, was in part responsible for the reading gains.

Seventh, the analysis of raters' judgments of overall quality shows that internal syntactic indices do not correlate with writing quality. Nold and Freedman (1977) and Faigley (1979b) had previously found Hunt's paradigm of syntactic maturity to be of very little value in predicting raters' judgments of quality. These two studies, however, did find the ratio of T-units with final nonrestrictive modifiers to be a weak influence on raters' judgments (about 16% in the Faigley study), but an influence nonetheless. The present study found final nonrestrictive modifiers to be of no more value than Hunt's indices. In light of these findings, the value of computing syntactic indices of any type in writing evaluation studies is questionable.

8.3 The Limitations of the Present Study

Although the present study was designed with a greater awareness of the many curricular and instructional variables which enter into the comparative evaluation of freshman writing courses than most previous studies, the present study is itself not without its limitations. Principal among the limitations is the size of the sample chosen from each of the twenty E-306 classes studied. In the case of all comparisons, whether "principal" or "derivative," reported in Chapter 4 and Chapter 5, the data from at least two classes were pooled, thus creating several comparisons in which an "n" of only 18 prevailed on one or both sides of the comparisons. We believe that such a small "n" may have precluded finding a larger number of significant effects for either instructional or curricular variables. Another major limitation of the present study is that the "synthetic-tutorial" classes wrote considerably less than did the students enrolled in other classes. This feature allowed the teachers of the "synthetic-tutorial" classes to spend a great deal of time guiding students in careful revisions of the essays they did write. In short, for a given essay the students in the "synthetic-tutorial" classes devoted more "time on task" to a particular writing assignment during the regular semester, while the students in the other classes--students who wrote a larger number of essays during the semester--spent considerably less time on any given assignment. This difference in instructional methodology represents a major uncontrolled variable in the present study, even though our results suggest that the instructional methodology employed in the "synthetic-tutorial" classes provides a suitable alternative for teaching composition. Another weakness is the failure of the present study to address either the question of teacher effectiveness or the question of student attitudes toward the study of writing in the various instructional settings investigated. Teacher evaluation data were collected from the students in the present study. These data were based on student responses to teacher effectiveness instruments provided by the University of Texas, but the questions asked were of such a general nature that they did not distinguish among teachers teaching different curricula or using different instructional methods. According to these data, each instructional method and each curricula was as effective as any other, was as liked as any other. Unless teacher effectiveness instruments can distinguish among teachers on the basis of instructional behaviors and techniques, they are of little use in comparative evaluations of the kind reported in the present study. A final major limitation of the present study stems from what we perceive as an overzealous control of the teacher variable in the present study. As we pointed out in Chapter 2 when we introduced the design of the five principal comparisons, a single teacher always taught two classes in any given comparison, with one class being on one side of the comparison and the other class being on the other side of the comparison. This control was thought initially to be a strength, but we suspect that in some cases, the unsuitability of a given teacher for a given composition curriculum or a given instructional method may have resulted in findings less reliable than we would have hoped. For example, in the comparison of the "synthetic-conventional" classes with the "analytic-conventional" classes, we strongly suspect that one of the two teachers was more comfortable with the "synthetic" curriculum while the other was more comfortable with the "analytic" curriculum. Both, however, were required by our research design to teach both curricula. The effect of this design demand was that the less-than-desirable performances of the two teachers, each in teaching a different curriculum, cancelled out the good performances they registered in teaching the curricula they were most comfortable in teaching.

In spite of these limitations, the present study still stands as perhaps the most carefully designed and carefully conducted comparative evaluation of two freshman writing courses that has even been completed. At the same time, it demonstrates that a great deal more work must be done before the profession at large can take much confidence in studies such as this one. Although the present study does much to erect an adequate theoretical framework for conducting comparative evaluations of college writing courses, that framework requires a great deal more thought than it has to date been given, either in the present study or in other studies. Among the other areas which need to be accommodated in such a framework are guidelines for assessing the goals of college writing programs in light of the demands made on students in other college courses and in light of the demands made on writers in the "real" world beyond the walls of the university or college. Such a framework should also better accommodate writing as a performative act than the present study does. Unfortunately, the present study, like all of its predecessors, addressed only the products the students produced, ignoring almost totally the acts which led to those products. It is incorrect to assume that if a curriculum does not immediately effect change in the products the students write, it is an inadequate curriculum. It is more realistic to assume that at least some of the effects may not, in fact, become apparent in the products students write until some months--perhaps even years--have passed after the student completed the composition course. Skill in using the written language develops slowly, and that fact must always be seen as a caveat in any study reporting significant or non-significant changes in the products student writers produce in connection with a given instructional treatment.

8.4 Implications for Teaching and Research

Any large and comprehensive study is bound to have something to say both to teachers of college freshmen and to those who are principally interested in addressing through empirical or theoretical research the ways that writing skills can best be learned and taught. The present study is, we believe, no exception.

Among the implications of the present study for the teaching of freshman composition are the following. First, the present study suggests that teachers may confidently use the generative rhetorical paradigm to teach argumentative writing to college freshmen. Indeed, as Harris and Witte (in press) point out, the potential of sentence-based approaches to freshman composition has not been realized; and the present study strongly suggests that "synthetic," sentence-based curricula can be used successfully to provide the instruction in rhetorical principles that college freshmen need to learn in order to write effectively regardless of discourse type.

The important issues are what specific rhetorical principles ought to be taught and how best to align those principles with composition course goals. This matter of focus, we believe, is crucial to the design of composition curricula and instruction. Composition courses which are not clearly focused are not likely to have the positive effects on student performance that were reported in the present study for the two curricula or that were reported for the "experimental" classes in the Miami University sentence-combining study or in the University of North Dakota study of generative rhetoric. We believe that both the courses in the present study and those used as "experimental" classes in the two previous studies were clearly focused courses, courses which had clearly defined sets of goals or objectives that could be articulated to the students enrolled in them. In contrast, most "traditional" freshman writing classes are not as clearly focused, relying as they do on textbooks such as McCrimmon's Writing with a Purpose which try to provide a little something for every teacher. While such a textbook publishing philosophy may insure greater sales, it does not seem to us that it has a happy effect on student writing, unless the textbook itself is made to serve the needs of a clearly focused course syllabus. Adelstein and Pival's The Writing Commitment--the textbook which served as the rhetoric in the "analytic" or "traditional" writing classes in the present study--is a textbook which, like Writing with a Purpose, was apparently published with the idea of providing something for every teacher. In the present study, however, the textbook was made clearly subordinate to a very carefully designed and clearly focused course syllabus reflecting an "aims and modes" approach to freshman composition. Both the students enrolled in the "analytic" or "traditional" courses examined in the present study and the teachers teaching those classes had a very clear sense of what the goals of the course were. Every assignment was specifically directed toward realizing one or more of those goals. We suspect that the lack of clearly articulated goals for "traditional" writing courses is the principal reason those courses seem always to come up short when compared to "sentence-based" approaches to freshman composition.

The present study also seems to have a number of implications for research in written composition. Paramount is the need for better design in writing program evaluation studies. This point has been made again and again in connection with the multitude of variables which influence the results of evaluation studies. Such studies require sophisticated designs employing a variety of assessment measures.

Several specific research issues are also raised. First, it seems to us that another major study of generative rhetoric is in order. Building on the previous work of Faigley at North Dakota and on the findings of the present study, a new major study of generative rhetoric could focus more sharply on the instructional and curricular variables conflated in generative rhetoric and could attempt to identify those variables which contribute to improved writing having different purposes and written in different modes. Second, our comparison of the design and the results of the present study to the design and the results of the Miami University sentence-combining study strongly points to the need for another comparison of the sentence-combining approach and the "traditional" approach to freshman composition. With well-designed courses on both sides of such a comparison and with adequate control of the major curricular and instructional variables, such as a study would do much to advance current pedagogical practice. Third, the relatively poor predictive power of the syntactic measures used in the present study strongly suggests the need to develop other indices for quantifying growth among college writers. On the basis of the detailed analyses in Chapter 3 of a small sample of the essays collected for the present study, we would speculate that such indices would have to measure, and thereby quantify, semantic relationships that extend across sentence boundaries instead of focusing on sentence-level structures or features. Fourth, our finding that neither of the curricula nor any of the instructional methodologies we studied seemed to affect the degree to which student writers were apprehensive about the writing they did indicates that composition researchers need (1) to develop specific ways to reducing writing apprehension among student writers and (2) to develop specific instructional methods for teaching student writers who are highly apprehensive. A fifth area which we see as potentially important would focus on the relationship between the ways students comprehend written texts they read and the decisions they make about the texts they write. Although reading and writing are two clearly distinct and distinguishable cognitive activities, it would seem profitable to determine whether the texts students write contain the same kinds of semantic structures they can easily process in reading.

The present study, of course, has other implications for both teaching and research. The few we have chosen to single out should suggest the importance of the present study in its larger context. The limitations and the implications of empirical studies are inextricably bound together, and the questions related to any issue of language use or development are complex and great in number. All that empirical researchers and composition teachers can expect is that each subsequent research study will answer some questions not answered in its predecessors. A few of these questions the present study answers. A great many more remain.

FIGURES AND TABLES

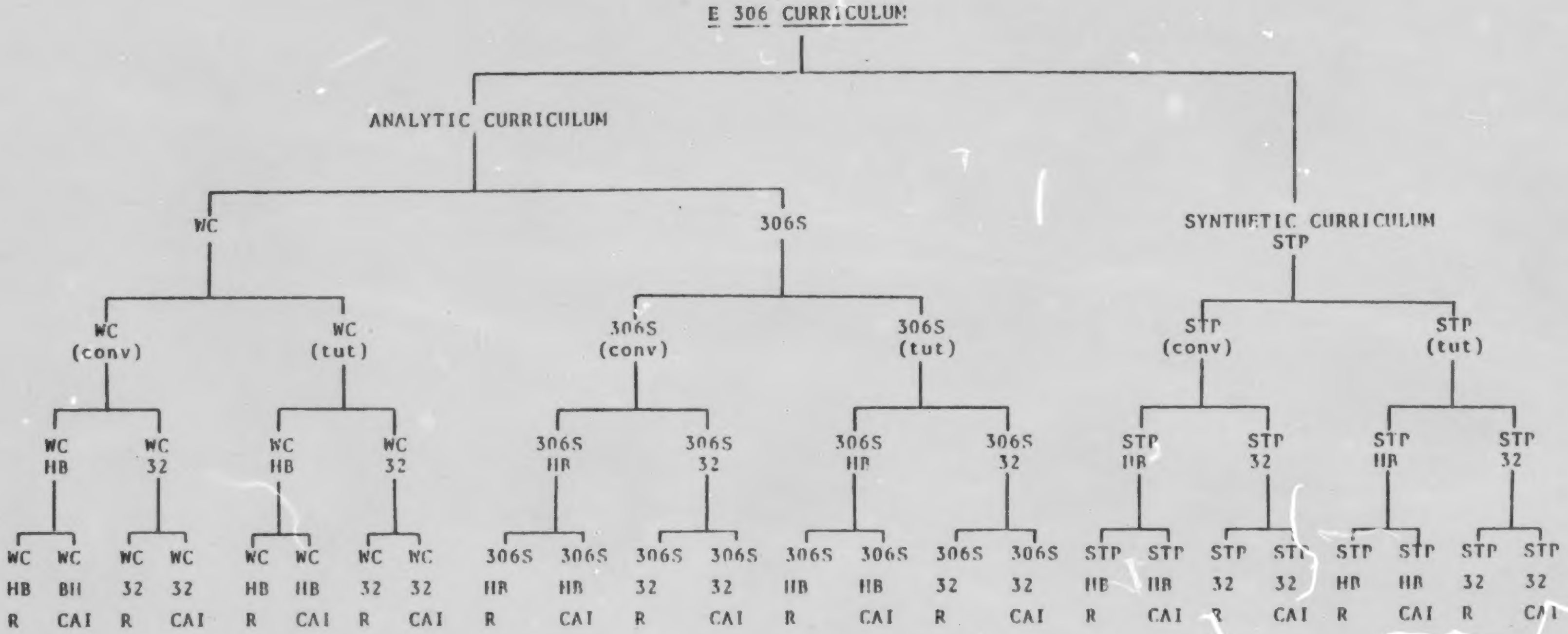
E 306 CURRICULUM
ANALYTIC CURRICULUM
WC
306S
SYNTHETIC CURRICULUM
STP
WC (conv)
WC (tut)
306S (conv)
306S (tut)
STP (conv)
STP (tut)
WC HB
WC 32
WC HB
WC 32
306S HB
306S 32
306S HB
306S 32
STP HB
STP 32
STP HB
STP 32
WC HB R
WC BH CAI
WC 32 R
WC 32 CAI
WC HB R
WC HB CAI
WC 32 R
WC 32 CAI
306S HB R
306S HB CAI
306S 32 R
306S 32 CAI
306S HB R
306S HB CAI
306S 32 R
306S 32 CAI
STP HB R
STP HB CAI
STP 32 R
STP 32 CAI
STP HB R
STP HB CAI
STP 32 R
STP 32 CAI

Figure 2.1. A Tree Diagram Illustrating the Possible Combinations of E306 Curricular and Instructional Variables. (The two E306 options described briefly in Chapter 1 are marked with asterisks; a legend appears on the following page.)

Legend for Figure 2.1

WC	Adelstein and Pival, *The Writing Commitment* (rhetoric text)
306S	Rodi, *Teaching Syllabus for E306* (rhetoric text)
STP	Wittig, *Steps to Structure* (rhetoric text)
conv	conventional (lecture/discussion) classroom instruction
tut	tutorial instruction
HB	Corder, *Handbook of Current English* (usage)
32	Blumenthal, *English 3200* (usage)
R	Decker, *Patterns in Exposition 6* (reader)
CAI	computer assisted instruction

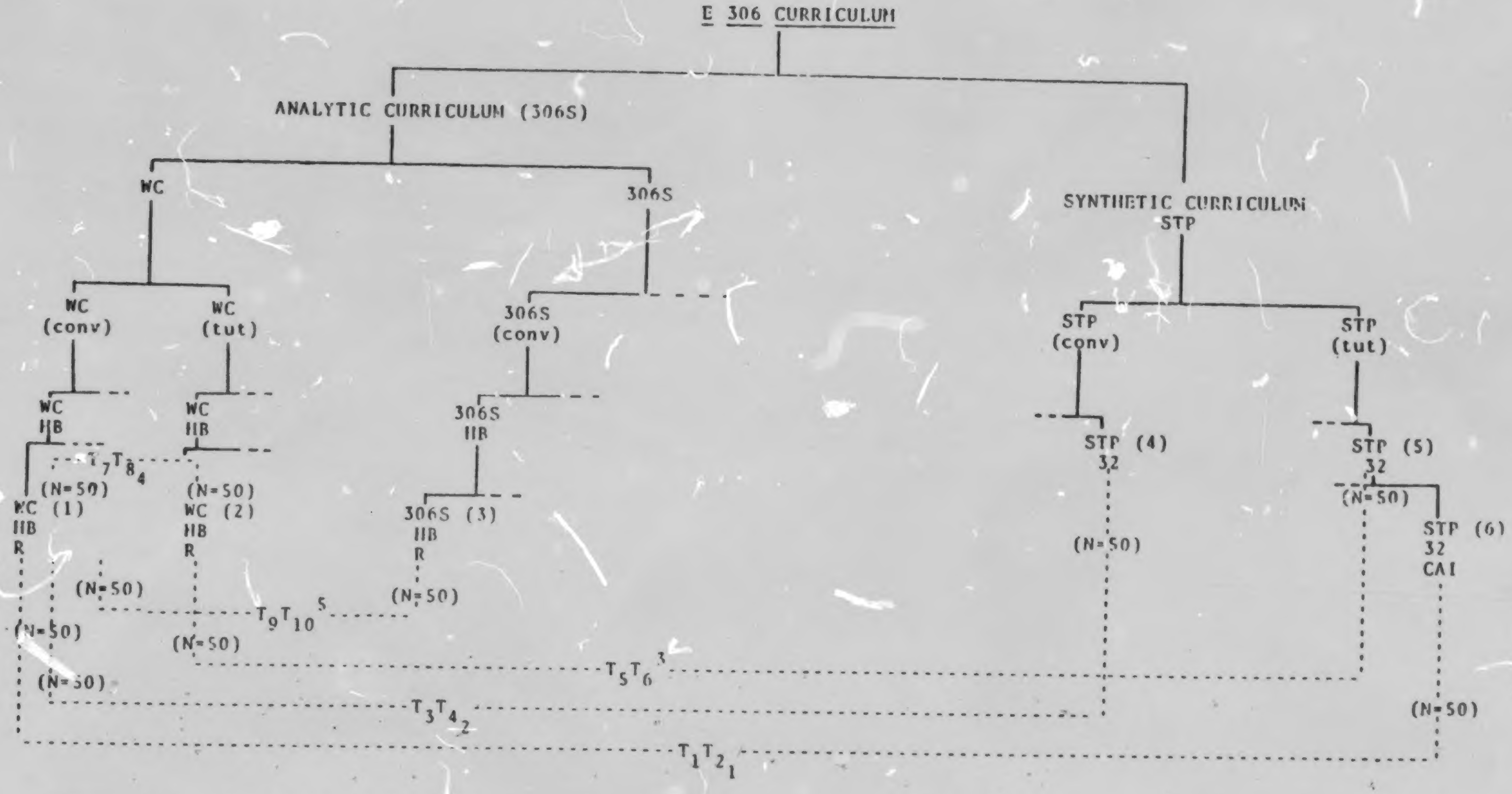
E 306 CURRICULUM
ANALYTIC CURRICULUM (306S)
SYNTHETIC CURRICULUM
STP
WC
306S
WC (conv)
WC (tut)
306S (conv)
STP (conv)
STP (tut)
WC HB
WC HB
306S HB
STP (4) 32
STP (5) 32
T7T8 4
(N=50)
WC (1) HB R
(N=50)
WC (2) HB R
306S (3) HB R
(N=50)
STP (6) 32 CAI
(N=50)
(N=50)
(N=50)
T9T10 5
(N=50)
(N=50)
(N=50)
T5T6 3
(N=50)
T3T4 2
(N=50)
T1T2 1

Figure 2.2. An Abridged Tree Diagram Showing the Five Principal Comparisons in the Present Study (A legend appears on the following page.)

Legend for Figure 2.2

WC	The Writing Commitment (rhetoric)
306S	Teaching Syllabus for E306 (rhetoric)
STP	Steps to Structure (rhetoric)
conv	Conventional Medium of Instruction (Classroom)
tut	Tutorial Medium of Instruction
HB	Handbook of Current English (Grammar/Usage)
32	English 3200 (Grammar/Usage)
R	Patterns of Exposition (Reader)
CAI	Computer-Assisted Instruction
T_1-T_{10}	Teachers (Each T designated for a given pair of courses were assigned one section [n=25] of each course in the pair.)
- - - -	Lines consisting of dashes connect courses paired for comparison.
(1-6)	Designates course compared.
1	Comparison of analytic-conventional course (1) with synthetic-tutorial course (6) -- testing for course effectiveness
2	Comparison of analytic-conventional course (1) with synthetic-conventional course (4) -- testing for effect of content
3	Comparison of analytic-tutorial course (2) with synthetic-tutorial course (5) -- testing for effect of content
4	Comparison of analytic-conventional course (1) with analytic-tutorial course (2) -- testing for effect of instructional medium
5	Comparison of analytic-conventional course (1) with analytic-conventional course (3) -- testing for effect of rhetoric text

Principal Comparisons	Section No. & Descriptor	A1/B1 Writing Topics				A2/B2 Writing Topics			
		Pretest		Posttest		Pretest		Posttest	
		A_1	B_1	A_1	B_1	A_2	B_2	A_2	B_2
1	01 AC/Conv	X			X		X	X	
	02 AC/Conv		X	X		X			X
	12 SC/LabTut		X	X		X			X
	04 SC/LabTut	X			X		X	X	
2	05 AC/Conv	X			X		X	X	
	06 AC/Conv		X	X		X			X
	07 SC/Conv		X	X		X			X
	08 SC/Conv	X			X		X	X	
3	17 AC/Conv	X			X		X	X	
	18 AC/Conv		X	X		X			X
	19 ACSyll/Conv		X	X		X			X
	20 ACSyll/Conv	X			X		X	X	
4	13 AC/Conv	X			X		X	X	
	14 AC/Conv		X	X		X			X
	15 AC/Tut		X	X		X			X
	16 AC/Tut	X			X		X	X	
5	09 AC/Tut	X			X		X	X	
	10 AC/Tut		X	X		X			X
	21 SC/Tut		X	X		X			X
	03 SC/Tut	X			X		X	X	

Table 2.1. Pretest and Posttest Administration of Writing Topics to Eliminate Bias ("X" indicates the pretest and posttest topics administered to each section).

Derivative Comparisons	N	Sections Involved
(1) AC/Conv	90	01, 02, 05, 06, 13, 14, 17, 18, 19, 20
SC/Conv	18	07, 08
(2) AC/Conv	90	01, 02, 05, 06, 13, 14, 17, 18, 19, 20
SC/Tut	36	12, 04, 21, 03
(3) Conv	108	01, 02, 05, 06, 07, 08, 13, 14, 17, 18, 19, 20
Tutorial	72	12, 04, 09, 10, 21, 03, 15, 16
(4) AC	126	01, 02, 05, 06, 09, 10, 13, 14, 15, 16, 17, 18, 19, 20
SC	54	12, 04, 07, 08, 21, 03
(5) AC/Conv	90	01, 02, 05, 06, 13, 14, 17, 18, 19, 20
AC/Tut	36	09, 10, 15, 16

Table 2.2. A Delineation of E306 Sections Involved in the Five Derivative Comparisons.

Scores Tests	1	2	3	4
A_1B_1 Pretests (N=360)	75	146	109	30
A_1B_1 Posttests (N=360)	54	128	119	59
A_2B_2 Pretests (N=360)	96	130	89	45
A_2B_2 Posttests (N=360)	66	100	108	86
A_1B_1 Pre- and Posttests (N=720)	129	274	228	89
A_2B_2 Pre- and Posttests (N=720)	162	230	197	131
A_1B_1 and A_2B_2 Pretests (N=720)	171	276	198	75
A_1B_1 and A_2B_2 Posttests (N=720)	120	228	227	145
All Topic, Pre and Post (N=1440)	291	504	425	220

Table 2.3. Distribution of Holistic Scores from Low Quality (1) to High Quality (4) for 720 E 306 Essays (N = no. of essays X no. of raters for each essay).

Tests \ Scores	1	2	3	4
A_1B_1 Pretests (N=360)	20.8	40.6	30.3	8.3
A_1B_1 Posttests (N=360)	15.0	35.6	33.0	16.4
A_2B_2 Pretests (N=360)	26.7	36.1	24.7	12.5
A_2B_2 Posttests (N=360)	18.3	27.8	30.0	23.9
A_1B_1 Pre- and Posttests (N=720)	17.9	38.0	31.7	12.4
A_2B_2 Pre- and Posttests (N-720)	22.5	31.9	27.4	18.2
A_1B_1 and A_2B_2 Pretests (N=720)	23.8	38.8	27.5	10.4
A_1B_1 and A_2B_2 Posttests (N=720)	16.7	31.7	31.5	20.1
All Topics, Pre and Post (N=1440)	20.2	35.0	29.5	15.3

Table 2.4. Percentage of Holistic Scores Distributed Across All Scoring Categories from Low Quality (1) to High Quality (4) for 720 E 306 Essays (N = no. of essays X no. of raters for each essay).

Scores Tests	2	3	4	5	6	7	8
A_1B_1 Pretests (N=180)	23	29	30	57	14	24	3
A_1B_1 Posttest (N=180)	11	32	27	42	19	39	10
A_2B_2 Prestest (N=180)	26	44	23	40	13	23	11
A_2B_2 Posttests (N=180)	17	32	18	32	17	42	22
A_1B_1 Pre- & Posttests (N=360)	34	61	57	99	33	63	13
A_2B_2 Pre- & Posttests (N=360)	43	76	41	72	30	65	33
A_1B_1 & A_2B_2 Pretests (N=360)	49	73	53	97	27	47	14
A_1B_1 & A_2B_2 Posttests (N=360)	28	64	45	74	36	81	32
All Topics Pre and Post (N=720)	77	137	98	171	63	128	46

Table 2.5. Distribution of Summed Holistic Scores from Low Quality (2) to High Quality (8) for E306 Essays (N = no. of essays).

Scores / Tests	2	3	4	5	6	7	8
A_1B_1 Pretests (N=180)	12.7	16.1	16.7	31.7	7.8	13.3	1.
A_1B_1 Posttests (N=180)	6.1	17.8	15.0	23.3	10.5	21.7	5.
A_2B_2 Pretests (N=180)	14.4	24.4	12.8	22.2	7.2	12.8	6.
A_2B_2 Posttests (N=180)	9.4	17.8	10.0	17.8	9.4	23.3	12.
A_1B_1 Pre- & Posttests (N=360)	9.4	16.9	15.8	27.5	9.2	17.5	3.
A_2B_2 Pre- & Posttests (N=360)	11.9	21.1	11.4	20.0	8.3	18.1	9.
A_1B_1 & A_2B_2 Pretests (N=360)	13.6	20.3	14.7	26.9	7.5	13.1	3.
A_1B_1 & A_2B_2 Posttests (N=360)	7.8	17.8	12.5	20.6	10.0	22.5	8.
All Topics Pre & Post (N=720)	10.7	19.0	13.6	23.8	8.7	17.8	6.

Table 2.6. Percentage of Holistic Scores Distributed Across All Scoring Categories from Low Quality (2) to High Quality (8) for E306 Essays (N = no. of essays).

Length Factors	Means for "High Quality" Essays	Means for "Low Quality" Essays	Differences Between Means
Words	645.6	270.4	375.2
Sentences	35.8	16.0	19.8
T-Units	42.2	19.8	22.4
Total Clauses	69.6	36.2	33.4

Table 3.1. Length Factors Describing Five "High Quality" and Five "Low Quality" E 306 Essays Written on Topic A_1.

Types of Error	Frequency of Occurrence (in words)		Difference
	High Quality Essays	Low Quality Essays	
Punctuation	189.88	79.53	111.35
End-Stop	1614.00	193.14	1420.86
Within-Sentence	215.20	135.20	80.00
Grammar	403.50	169.00	234.50
Verb	1614.00	676.00	938.00
Pronoun	645.60	270.40	375.20
Modification	3228.00	1352.00	1876.00
Spelling	269.00	64.38	204.62
ALL ERRORS	87.24	29.39	57.85

Table 3.2. The Frequency (in words) of Certain Common Usage Errors of Three Types in "High" and "Low" Quality E 306 Essays.

Types of Error	Frequency of Occurrence (in T-Units)		Difference
	High Quality Essays	Low Quality Essays	
Punctuation	12.41	5.82	6.59
End-Stop	105.50	14.14	91.36
Within-Sentence	14.07	9.90	4.17
Grammar	26.38	12.38	14.00
Verb Tense	105.50	49.59	56.00
Pronoun	42.50	19.80	22.70
Modification	211.00	99.00	112.00
Spelling	17.58	4.71	12.87
ALL ERRORS	5.70	2.15	3.55

Table 3.3. The Frequency (in T-units) of Certain Common Usage Errors of Three Types in "High" and "Low" Quality E 360 Essays.

Types of Error	Frequencies of Error		
	in words	in T-units	per 1000 words
Punctuation	134.7	9.12	7.42
End-Stop	508.8	34.44	1.97
Within-sentence	183.2	12.40	5.46
Grammar	286.3	19.38	3.49
Verbs	1145.0	77.50	0.87
Pronouns	458.0	31.00	2.18
Modification	2290.0	155.00	0.44
Spelling	138.8	9.40	7.20
ALL ERRORS	55.2	3.73	18.12

Table 3.4. Frequency of Error Calculated in Terms of Total Words and Total T-units and Per 1000 Words in the Total Discourse Sample Collected from Five "High" and Five "Low" Quality E306 Students.

	Mean % of Total Words in Free Modifiers	% in Initial Free Modifiers	% in Medial Free Modifiers	% in Final Free Modifiers
University of N.D Freshmen	16.1	11.8	0.8	3.5
Hall-Emblen Writers	30.03	8.8	4.2	17.3

Table 3.5. Faigley's (1979a) Comparison of the Occurrence of Free Modifiers in the Writing of University of North Dakota Freshmen and of Writers Anthologized in Hall and Emblen's *A Writer's Reader*.

Indices of Syntactic Complexity	Means for "High" and "Low" Quality Essays	Means for "High" Quality Essays	Means for "Low" Quality Essays	Differences Between Means of "High and "Low" Quality Essays
words/T-units	14.48	15.30	13.66	1.64
words/clauses	8.33	9.28	7.47	1.81
clauses/T-units	1.74	1.65	1.83	-0.18
T-units/sentences	1.21	1.18	1.24	-0.06
words/sentences	17.48	18.03	19.90	1.13

Table 3.6. A Comparison of the Syntactic Complexity of "High" and "Low" Quality E 306 Essays, Based on Hunt's (1965) Five Indices of Syntactic Maturity or Complexity.

	Mean & of Total Words in Free Modifiers	% in Initial Position	% in Medial Position	% in Final Position
"High" Essays	10.56	4.12	1.61	4.83
"Low" Essays	5.77	-----	1.63	4.14
All Essays	9.15	2.90	1.62	4.63

Table 3.7. Percent of Total Words in Free Modifiers in "High" and "Low" Quality E 306 Essays.

	Mean % of Total T-units with Free Modifiers	% with Initial Free Modifiers	% with Medial Free Modifiers	% with Final Free Modifiers
"High Essays	28.44	13.74	6.64	8.06
"Low" Essays	10.10	-------	3.03	7.07
All Essays	22.58	9.35	5.48	7.74

Table 3.8. Percent of Total T-units with Free Modifiers in "High" and "Low" Quality E 306 Essays.

Indices and Studies	G3	G4	G5	G7	G8	G12	Coll. Fresh.	Skilled Adults
Words/T-units								
Cooper, et al.							16.54	20.46
Faigley							14.50	19.00
Hunt		8.60			11.50	14.40		20.30
O'Donnell, et al.	7.67		9.34					
Stewart							14.60	
Sodowsky & Witte							15.12	
Present							14.48	
Words/Clauses								
Cooper, et al.							10.16	11.60
Faigley							8.90	11.30
Hunt		6.60			8.10	8.60		11.50
O'Donnell, et al.	6.50		7.40	7.70				
Stewart							8.76	
Sodowsky & Witte							9.28	
Present							8.33	
Clauses/T-units								
Cooper, et al.								
Faigley							1.62	1.67
Hunt		1.30			1.42	1.68		1.78
O'Donnell, et al.	1.18		1.27	1.30				
Stewart								
Sodowsky & Witte							1.66	
Present							1.74	
T-units/Sentences								
Cooper, et al.								
Faigley								
Hunt		1.60			1.37	1.17		1.23
O'Donnell, et al.								
Stewart								
Sodowsky & Witte							1.18	
Present							1.21	
Words/Sentences								
Cooper, et al.								
Faigley								
Hunt		13.50			15.90	16.90		24.70
O'Donnell, et al.								
Stewart								
Sodowsky & Witte							17.57	
Present							17.48	

Table 3.9. The Pooled Scores for "High" and "Low" Quality E306 Essays ("Present") for Hunt's Five Indices in Comparison to Scores Reported by Cooper, et al. (1979), FAigley (1979b), Hunt (1965), O'Donnell, et al. (1967), Stewart (1978a), and Sodowsky and Witte (1978).

	% of Words in Medial Free Mods.	% of Words in Final Free Mods.	% of T-units with Final Free Mods.
E 306 Essays	1.62	4.63	7.74
U of N. Dakota Essays	N.A.	3.81	5.40
Stanford U. Essays	2.00	3.00	N.A.

Table 3.10. A Comparison of the Occurrence of Free Modifiers in the Written Discourse of E 306 Essays, University of North Dakota Essays (Faigley, 1979b), and Stanford University Essays (Nold and Freedman, 1977).

Pronoun Type	% of Words in "High" Quality Essays	% of Words in "Low" Quality Essays	Difference
First Person	0.49	0.89	-0.40
Second Person	0.06	3.55	-3.49
First & Second Person	0.56	4.44	-3.88

Table 3.11. The Occurrence of First- and Second-Person Pronouns Expressed in Terms of the Percentage of Words in "High" and "Low" Quality E 306 Essays.

Pronoun Type	Frequency in "High" Quality Essays		Frequency in "Low" Quality Essays		Differences	
	in words	in T-units	in words	in T-units	in words	in T-units
First Person	201.75	13.19	112.67	8.25	89.08	4.94
Second Person	1614.00	105.50	28.17	2.06	1585.83	103.44
First & Second Person	179.33	11.72	22.53	1.65	156.80	10.07

Table 3.12. The Frequency in Words and T-Units of the Occurrence of First- and Second Person Pronouns in "High" and "Low" Quality E 306 Essays.

	W / CT Type		CT Type / T		% of CT		CT per 100 T	
	Low	High	Low	High	Low	High	Low	High
Per Pronoun	233.00	63.40	0.510	0.257	2.10	5.04	5.10	25.70
Demon. & Art.	77.67	58.89	0.156	0.277	6.29	5.44	15.60	27.70
Comparative	43.69	53.19	0.271	0.307	11.18	6.02	27.10	30.70
REFERENCE (all)	24.96	19.40	0.478	0.841	19.57	16.50	47.80	84.10

Table 3.13. Summary of Reference Cohesion in "High" and "Low" Quality E 306 Essays.

	W / CT Type		CT Type / T		% of CT		CT per 100 T	
	Low	High	Low	High	Low	High	Low	High
Nominal Sub.	349.5	329.8	0.034	0.050	1.40	0.96	3.40	5.00
Verbal Sub.	---	329.8	---	0.050	---	0.96	---	5.00
Clausal Sub.	---	549.7	---	0.029	---	0.57	---	2.90
SUBSTITUTION (All)	349.5	126.9	0.034	0.129	1.40	2.49	3.40	12.90

Table 3.14. Summary of Substitution Cohesion in "High" and "Low" Quality E 306 Essays.

	W / CT Type		CT Type / T		% of CT		CT per 100 T	
	Low	High	Low	High	Low	High	Low	High
Nominal Ell.	139.8	329.8	0.085	0.049	3.50	0.96	8.50	5.00
Verbal Ell.	699.0	824.5	0.017	0.020	0.70	0.38	1.70	1.90
Clausal Ell.	---	235.6	---	0.069	---	1.34	---	6.90
ELLIPSIS (all)	116.5	117.8	0.102	0.138	4.20	2.68	10.20	13.80

Table 3.15. Summary of Ellipsis Cohesion in "High" and "Low" Quality E 306 Essays.

	W / CT Type		CT Type / T		% of CT		CT per 100 T	
	Low	High	Low	High	Low	High	Low	High
additive conj.	116.50	71.70	0.102	0.288	4.20	4.41	10.20	22.80
adversative conj.	---	164.90	---	0.099	---	1.91	---	9.90
causal conj.	699.00	117.79	0.017	0.139	0.70	2.68	1.70	13.90
temporal conj.	139.80	206.13	0.085	0.079	3.50	1.53	8.50	7.90
continuative conj.	---	149.91	---	0.109	---	2.11	---	10.90
CONJUNCTION (all)	58.25	24.99	0.204	0.654	8.40	12.64	20.40	65.40

Table 3.16. Summary of Conjunctive Cohesion in "High" and "Low" Quality E 306 Essays.

	W / CT Type		CT Type / T		% of CT		CT per 100 T	
	Low	High	Low	High	Low	High	Low	High
same item	11.27	9.11	1.051	1.792	43.35	34.67	105.10	179.20
synonym	58.25	33.65	0.203	0.485	8.39	9.39	20.30	48.50
superordinate	---	274.83	---	0.059	---	1.15	---	5.90
general item	174.75	126.85	0.068	0.129	3.80	2.49	6.80	12.90
collocation	41.12	17.36	0.288	0.941	11.89	18.20	28.80	94.10
LEXICAL (all)	7.36	4.79	1.610	3.406	66.43	65.90	161.00	340.60

Table 3.17. Summary of Lexical Cohesion in "High" and "Low" Quality E 306 Essays.

	"Low" Quality Essays	"High" Quality Essays
same item	65.26 %	52.62 %
synonym	12.63 %	14.24 %
superordinate	---	1.74 %
general item	4.21 %	3.78 %
collocation	17.90 %	27.62 %

Table 3.18. Percentage of Lexical Ties Attributed to the Five Types of Lexical Cohesion in "High" and "Low" Quality E 306 Essays.

	Ratio of Ties to Words		Cohesive Type as % of Total Words	
	Low	High	Low	High
Reference	0.0400	0.0515	4.00	5.15
Substitution	0.0029	0.0080	0.29	0.80
Ellipsis	0.0086	0.0085	0.86	0.85
Conjunction	0.0172	0.0400	1.72	4.00
Lexical	0.1359	0.2086	13.59	20.86
ALL COHESIVE TIES	0.2046	0.3166	20.46	31.66

Table 3.19. Summary of the Ratio of Cohesive Ties to Total Words (Together with a Conversion to Percentages of Total Words) for "High" and "Low" Quality E 306 Essays.

	"Low" Quality Essays	"High" Quality Essays
Immediate Ties	32.80%	41.59%
Remote Ties	26.69%	26.94%
Mediated Ties	0.79%	5.61%
Remote-Mediated Ties	36.72%	25.86%

Table 3.20. Percentage of Total Cohesive Ties Represented by Four "Span of Text" Cohesive Ties Categories in "High" and "Low" Quality E 306 Essays.

	Pretest		Posttest					
Factor	Mean	S.D	Mean	S.D	Difference	S.D	T-value	Significance
A_1B_1 Topics								
Words	429.35	148.90	429.98	136.20	-0.63	117.58	-0.07	N.S.
Words/T-units	15.26	3.02	15.28	2.77	-0.02	3.46	-0.08	N.S.
Words/Clauses	8.92	1.51	9.03	1.47	-0.11	1.84	-0.81	N.S.
Ratio of TUs with FFM	0.048	0.06	0.050	0.06	-0.002	0.07	-0.36	N.S.
Ratio of Words in FFM	0.027	0.04	0.029	0.04	-0.002	0.05	-0.43	N.S.
Holistic Score	4.52	1.58	5.02	1.70	-0.49	2.02	-3.28	.001
A_2B_2 Topics								
Words	438.47	134.99	477.40	144.22	-38.93	132.24	-3.95	.001
Words/T-units	16.96	3.61	16.84	3.38	0.12	4.27	0.38	N.S.
Words/Clauses	10.54	1.65	10.40	1.85	0.14	2.15	0.89	N.S.
Ratio of TUs with FFM	0.039	0.05	0.048	0.06	-0.009	0.07	-1.72	N.S.
Ratio of Words in FFM	0.020	0.03	0.027	0.03	-0.006	0.04	-2.10	.037
Holistic Score	4.47	1.80	5.19	1.94	-0.72	2.35	-4.13	.001

Table 4.1. Paired T-Tests (N=180) of Change Scores on Writing Samples for All 180 Students in the 20 E 306 Sections Regardless of Instructional or Curricular Treatment.

Indices and Studies	G3	G4	G5	G7	G8	G12	College Freshmen	Skilled Adults
Words/T-units								
Cooper, et al.							16.54	20.46
Faigley							14.50	19.00
Hunt		8.60			11.50	14.40		20.30
O'Donnell, et al.	7.67		9.34	9.77				
Stewart							14.60	
Sodowsky & Witte							15.12	
A_1B_1 Topics							15.26	
A_2B_2 Topics							16.96	
Words/Clauses								
Cooper, et al.							10.16	11.60
Faigley							8.90	11.30
Hunt		6.60			8.10	8.60		11.50
O'Donnell, et al.	6.50		7.40	7.70				
Stewart							8.76	
Sodowsky & Witte							9.29	
A_1B_1 Topics							8.92	
A_2B_2 Topics							10.54	

Table 4.2. The Means for A_1B_1 and A_2B_2 E 306 Pretest Essays for "Words per T-Unit" and "Words per Clause" in Comparison to Means Reported by Cooper, et al. (1980), Faigley (1979b), Hunt (1965), O'Donnell, et al. (1967), Stewart (1978a), and Sodowsky and Witte (1978).

McGraw-Hill Writing Test: Forms and Sections	Mean	S.D	Difference Between Mean Scores
Section I: Language Mechanics			
Form A	17.49	4.40	2.86
Form B	20.35	4.50	
Section II: Sentence Patterns			
Form A	15.78	3.95	0.34
Form B	15.44	4.40	
Section III: Paragraph Patterns			
Form A	10.69	2.69	0.63
Form B	10.06	2.68	
Total Raw Score			
Form A	43.96	N.A.	1.90
Form B	45.86	N.A.	

Table 4.3. Means for the McGraw-Hill Writing Test Norming Group (see Raygor, 1970b, p. 26).

McGraw-Hill Reading Test: Paragraph Comprehension	Mean	S.D	Difference Between Mean Scores
Form A	18.57	4.84	0.64
Form B	19.21	4.90	

Table 4.4. Means for the McGraw-Hill Reading Test: Paragraph Comprehension Norming Group (see Raygor, 1970a, p. 30).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	165	18.86	3.36	22.01	4.02	-3.15	3.38	-10.43	.001
Sentence Patterns	165	16.22	3.19	17.04	3.31	-0.82	3.33	- 3.16	.002
Paragraph Patterns	165	11.64	2.43	11.39	1.76	0.25	2.77	1.15	N.S.
Total Writing	165	46.72	6.36	50.44	7.12	-3.72	6.27	- 7.63	.001
McGraw-Hill Reading Test									
Reading Comprehension	159	20.11	3.99	20.89	3.56	-0.78	4.26	- 2.33	.021
Daly-Miller Writing Apprehension Test	159	69.92	16.09	70.37	16.18	-0.45	11.21	- 0.50	N.S.

Table 4.5. Paired T-Tests of Change Scores on Objective Tests for Students in the 20 E 306 Sections Regardless of Instructional or Curricular Treatment.

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	441.43	146.63	421.63	129.04	19.80	122.62	1.53	N.S.
Words/T-units	15.41	3.05	15.29	2.87	0.12	3.54	0.31	N.S.
Words/Clauses	8.94	1.46	9.10	1.65	- 0.16	2.14	-0.73	N.S.
Ratio of TUs with FFM	0.051	0.05	0.056	0.07	- 0.005	0.08	-0.63	N.S.
Ratio of Words in FFM	0.030	0.04	0.033	0.04	- 0.003	0.05	-0.58	N.S.
Holistic Score	4.53	1.52	5.19	1.61	- 0.66	1.95	-3.19	.002
A_2B_2 Topics								
Words	442.89	142.66	454.36	130.62	-11.47	116.86	-0.93	N.S.
Words/T-units	17.05	3.80	16.82	3.63	0.23	4.69	0.47	N.S.
Words/Clauses	10.62	1.73	10.25	1.85	0.36	2.38	1.44	N.S.
Ratio of TUs with FFM	0.042	0.05	0.038	0.05	0.004	0.07	0.59	N.S.
Ratio of Words in FFM	0.022	0.03	0.020	0.03	0.002	0.04	0.51	N.S.
Holistic Score	4.73	1.87	5.37	1.95	- 0.63	2.32	-2.60	.011

Table 4.6. Paired T-Tests (N=90) of Change Scores on Writing Samples for All "Analytic-Conventional" Sections (01, 02, 05, 06, 13, 14, 17, 18, 19, and 20) in Derivative Comparisons 1, 2, and 5 (see Table 2.2).

Factor	Pretest Mean	S.D.	Posttest Mean	S.D	Difference	S.D.	T-Value	Significance
A$_1$B$_1$ Topics								
Words	465.00	160.29	480.39	152.42	- 15.39	82.95	-0.79	N.S.
Words/T-units	14.43	2.80	15.54	3.00	- 1.11	3.13	-1.50	N.S.
Words/Clauses	8.70	1.53	9.18	1.25	- 0.49	1.19	-1.75	N.S.
Ratio of TUs with FFM	0.020	0.02	0.031	0.03	- 0.012	0.04	-1.21	N.S.
Ratio of Words in FFM	0.011	0.02	0.020	0.02	- 0.008	0.03	-1.29	N.S.
Holistic Score	4.44	1.29	5.11	1.91	- 0.67	2.50	-1.13	N.S.
A$_2$B$_2$ Topics								
Words	458.94	152.18	575.94	189.72	-117.00	125.87	-3.94	.001
Words/T-units	15.86	3.20	17.50	3.70	- 1.64	3.51	-1.98	N.S.
Words/Clauses	10.57	1.89	10.74	1.38	- 0.17	1.37	-0.53	N.S.
Ratio of TUs with FFM	0.032	0.04	0.064	0.05	- 0.032	0.06	-2.38	.029
Ratio of Words in FFM	0.018	0.03	0.037	0.03	- 0.019	0.04	-2.22	.041
Holistic Score	4.33	1.72	5.06	1.83	- 0.72	2.42	-1.27	N.S.

Table 4.7. Paired T-tests (N=18) of Change Scores on Writing Samples for "Synthetic-Conventional" Sections (07, 08) in Derivative Comparison 1 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	2.98	N.S.
T-Unit Length	0.55	N.S.
Clause Length	0.08	N.S.
Ratio of T-Units with Final Free Modifiers	1.12	N.S.
Ratio of Words in Final Free Modifiers	1.01	N.S.
Holistic Score	0.21	N.S.

Table 4.8. Comparison of Posttest Means on A_1B_1 Topic Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections in Derivative Comparison 1 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	16.08	.001
T-Unit Length	1.08	N.S.
Clause Length	1.19	N.S.
Ratio of T-Units with Final Free Modifiers	4.37	.039
Ratio of Words in Final Free Modifiers	4.93	.029
Holistic Score	0.19	N.S.

Table 4.9. Comparison of Posttest Mean Scores on A_2B_2 Topic Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections in Derivative Comparison No. 1 (see Table 2.2).

Factor	N	Pretest		Posttest		Difference	S.D.	T-Value	Significance
		Mean	S.D.	Mean	S.D				
McGraw-Hill Writing Test									
Language Mechanics	80	18.92	3.40	21.79	4.29	- 2.87	4.29	- 5.99	.001
Sentence Patterns	80	16.65	2.97	16.90	2.80	- 0.25	3.12	- 0.72	N.S.
Paragraph Patterns	80	11.45	2.85	11.43	1.90	0.03	3.36	0.07	N.S.
Total Writing	80	47.01	5.93	50.11	7.06	- 3.10	6.92	- 4.01	.001
McGraw-Hill Reading Test									
Reading Comprehension	82	20.28	3.24	21.46	2.85	- 1.18	3.57	- 3.00	.004
Daly-Miller Writing Apprehension Test	84	68.39	16.21	69.17	17.11	- 0.79	11.42	- 0.63	N.S.

Table 4.10. Paired T-Tests of Change Scores on Objective Tests for All "Analytic-Conventional" Sections (01, 02, 05, 06, 13, 14, 17, 18, 19, and 20) in Derivative Comparisons 1, 2, and 5 (see Table 2.2).

Factor	N	Pretest Mean	S.D.	Posttest Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	16	19.37	2.31	23.32	2.44	- 3.94	3.24	- 4.87	.001
Sentence Patterns	16	15.37	1.93	17.94	2.54	- 2.56	2.96	- 3.46	.004
Paragraph Patterns	16	12.50	1.27	11.19	1.11	1.31	1.92	2.73	.015
Total Writing	16	47.25	4.12	52.44	4.84	- 5.19	6.18	- 3.36	.004
McGraw-Hill Reading Test									
Reading Comprehension	17	20.94	4.15	20.53	3.48	0.41	4.54	0.37	N.S.
Daly-Miller Writing Apprehension Test	18	70.17	14.89	69.11	17.70	1.06	9.76	0.46	N.S.

Table 4.11. Paired T-Tests of Change Scores on Objective Tests for "Synthetic-Conventional" Sections (07, 08) in Derivative Comparison No. 1 (see Table 2.2).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Patterns	1.60	N.S.
Sentence Patterns	4.55	.036
Paragraph Patterns	0.25	N.S.
Total Score	1.70	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	2.18	N.S.
Daly-Miller Writing Apprehension Test	0.30	N.S.

Table 4.12. Comparison of Posttest Mean Scores on Objective Test Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections in Derivative Comparison No. 1 (see Table 2.2).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	375.53	160.67	435.72	153.51	-60.19	103.59	- 3.49	.001
Words/T-units	15.97	2.85	15.35	2.38	0.62	3.26	1.15	N.S.
Words/Clauses	9.45	1.83	9.06	1.27	0.39	1.71	1.38	N.S.
Ratio of TUs with FFM	0.045	0.06	0.041	0.04	0.004	0.06	0.35	N.S.
Ratio of Words in FFM	0.024	0.04	0.025	0.03	- 0.001	0.04	- 0.10	N.S.
Holistic Score	3.97	1.77	4.44	1.75	- 0.47	2.10	- 1.35	N.S.
A_2B_2 Topics								
Words	428.33	123.23	500.44	151.03	-72.11	168.79	- 2.56	.015
Words/T-units	17.55	3.72	17.14	3.09	0.41	4.02	0.61	N.S.
Words/Clauses	10.50	1.65	10.80	2.37	0.30	2.44	- 0.74	N.S.
Ratio of TUs with FFM	0.036	0.04	0.073	0.07	- 0.037	0.08	- 2.68	.011
Ratio of Words in FFM	0.019	0.02	0.043	0.04	- 0.024	0.04	- 3.35	.002
Holistic Score	4.03	1.77	4.83	2.17	- 0.80	2.57	- 1.88	N.S.

Table 4.13. Paired T-Tests (N=36) of Change Scores on Writing Samples for All "Synthetic-Tutorial" Sections (12, 04, 21, 03) in Derivative Comparison 2 (see Table 2.2).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	33	18.36	4.16	20.85	4.53	- 2.49	3.35	- 4.27	.001
Sentence Patterns	33	15.40	3.82	16.88	4.66	- 1.48	3.22	- 2.65	.012
Paragraph Patterns	33	11.27	2.32	11.40	2.00	- 0.12	2.06	- 0.34	N.S.
Total Writing	33	45.03	8.44	49.12	9.47	- 4.09	5.57	- 4.22	.001
McGraw-Hill Reading Test									
Reading Comprehension	30	19.13	5.95	20.23	4.51	- 1.10	5.38	- 1.12	N.S.
Daly-Miller Writing Apprehension Test	27	71.03	17.43	71.74	15.63	- 0.70	13.16	- 0.28	N.S.

Table 4.14. Paired T-Tests of Change Scores on Objective Tests for All "Synthetic-Tutorial" Sections (12, 04, 21, and 03) in Derivative Comparison 2 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	6.95	.009
T-Unit Length	0.03	N.S.
Clause Length	0.15	N.S.
Ratio of T-Units with Final Free Modifiers	1.18	N.S.
Ratio of Words in Final Free Modifiers	0.81	N.S.
Holistic Score	3.51	N.S.

Table 4.15. Comparison of Posttest Mean Scores on A_1B_1 Topics Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Tutorial" Sections in Derivative Comparison 2 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	5.42	.022
T-Unit Length	0.10	N.S.
Clause Length	2.12	N.S.
Ratio of T-Units with Final Free Modifiers	9.74	.001
Ratio of Words in Final Free Modifiers	11.96	.001
Holistic Score	0.80	N.S.

Table 4.16. Comparison of Posttest Scores on A_2B_2 Topics Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Tutorial" Sections in Derivative Comparison 2 (see Table 2.2).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	0.60	N.S.
Sentence Patterns	1.43	N.S.
Paragraph Patterns	0.01	N.S.
Total Score	0.09	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	1.63	N.S.
Daly-Miller Writing Apprehension Test	0.05	N.S.

Table 4.17. Comparison of Posttest Mean Scores on Objective Test Data with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Tutorial" Sections in Derivative Comparison 2 (see Table 2.2).

Factor	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	445.36	148.47	431.43	134.25	13.93	117.36	1.23	N.S.
Words/T-units	15.25	3.02	15.33	2.88	- 0.08	3.49	-0.26	N.S.
Words/Clauses	8.90	1.47	9.12	1.58	- 0.22	2.01	-1.13	N.S.
Ratio of TUs with FFM	0.045	0.05	0.052	0.06	- 0.007	0.07	-0.89	N.S.
Ratio of Words in FFM	0.027	0.03	0.031	0.04	- 0.004	0.05	-0.85	N.S.
Holistic Score	4.52	1.48	5.18	1.66	- 0.66	2.04	-3.35	.001
A_2B_2 Topics								
Words	445.57	143.68	474.62	148.26	-29.06	124.25	-2.43	.017
Words/T-units	16.85	3.71	16.93	3.64	- 0.08	4.56	-0.18	N.S.
Words/Clauses	10.61	1.75	10.33	1.78	0.28	2.25	1.26	N.S.
Ratio of TUs with FFM	0.041	0.05	0.043	0.05	- 0.002	0.07	-0.30	N.S.
Ratio of Words with FFM	0.022	0.03	0.023	0.03	- 0.001	0.04	-0.38	N.S.
Holistic Score	4.67	1.84	5.31	1.93	- 0.64	2.32	-2.90	.005

Table 4.18. Paired T-Tests (N=108) of Change Scores on Writing Samples for All "Conventional" Section (01, 02, 05, 06, 07, 08, 13, 14, 17, 18, 19, and 20) in Derivative Comparison 3 (see Table 2.2).

Factor	Pretest		Posttest					
	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	405.33	147.31	427.81	139.99	-22.47	115.28	- 1.65	N.S.
Words/T-units	15.28	3.05	15.21	2.61	0.07	3.43	0.20	N.S.
Words/Clauses	8.96	1.57	8.91	1.29	0.05	1.55	0.28	N.S.
Ratio of TUs with FFM	0.051	0.06	0.047	0.06	0.004	0.07	0.52	N.S.
Ratio of Words in FFM	0.029	0.04	0.026	0.03	0.003	0.04	0.45	N.S.
Holistic Score	4.53	1.74	4.78	1.75	- 0.25	1.98	-1.07	N.S.
A_2B_2 Topics								
Words	427.83	121.00	481.57	138.86	-53.74	143.01	-3.19	.002
Words/T-units	17.13	3.46	16.71	2.97	0.42	3.81	-.93	N.S.
Words/Clauses	10.44	1.50	10.49	1.94	- 0.05	2.00	-0.22	N.S.
Ratio of TUs with FFM	0.037	0.05	0.057	0.06	- 0.020	0.08	-2.19	.032
Ratio of Words in FFM	0.018	0.03	0.032	0.04	- 0.014	0.04	-2.77	.007
Holistic Score	4.17	1.71	5.00	1.96	- 0.83	2.40	-2.95	.004

Table 4.19. Paired T-Tests (N=72) of Change Scores on Writing Samples for All "Tutorial" Sections (12, 04, 09, 10, 21, 03, 15, and 16) in Derivative Comparison 3 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	1.80	N.S.
T-Unit Length	0.12	N.S.
Clause Length	1.04	N.S.
Ratio of T-Units with Final Free Modifiers	0.51	N.S.
Ratio of Words in Final Free Modifiers	0.70	N.S.
Holistic Score	2.55	N.S.

Table 4.20. Comparison of Posttest Mean Scores on A_1B_1 Topics Data with Pretest Scores Covaried for "Conventional" and "Tutorial" Sections in Derivative Comparison 3 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	0.91	N.S.
T-Unit Length	0.34	N.S.
Clause Length	0.56	N.S.
Ratio of T-Units with Final Free Modifiers	2.99	N.S.
Ratio of Words in Final Free Modifiers	3.60	N.S.
Holistic Score	0.48	N.S.

Table 4.21. Comparison of Posttest Mean Scores on A_2B_2 Topics Data with Pretest Scores Covaried for "Conventional" and "Tutorial" Sections in Derivative Comparison 3 (see Table 2.2).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	96	18.99	3.24	22.05	4.07	- 3.05	4.19	- 7.23	.001
Sentence Patterns	96	16.44	2.85	17.07	2.78	- 0.64	3.20	- 2.02	.050
Paragraph Patterns	96	11.63	2.67	11.39	1.79	0.24	3.19	0.73	N.S.
Total Writing	96	47.05	5.65	50.50	6.77	- 3.45	6.81	- 4.96	.001
McGraw-Hill Reading Test									
Reading Comprehension	99	20.40	3.40	21.30	2.97	- 0.91	3.77	0 2.39	.019
Daly-Miller Writing Apprehension Test	102	68.71	15.93	69.17	17.13	- 0.46	11.13	- 0.42	N.S.

Table 4.22. Paired T-Tests of Change Scores on Objective Tests for All "Conventional" Sections (01, 02, 05, 06, 07, 08, 13, 14, 17, 18, 19, and 20) in Derivative Comparison 3 (see Table 2.2).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	69	18.68	3.54	21.99	3.97	- 3.29	3.52	- 7.77	.001
Sentence Patterns	69	15.91	3.60	16.99	3.96	- 1.07	3.50	- 2.55	.013
Paragraph Patterns	69	11.65	2.06	11.39	1.73	0.26	2.06	1.05	N.S.
Total Writing	69	42.65	7.26	50.35	7.62	- 4.10	5.44	- 6.26	.001
McGraw-Hill Reading Test									
Reading Comprehension	60	19.63	4.81	20.22	4.31	- 0.58	4.97	- 0.91	N.S.
Daly-Miller Writing Apprehension Test	57	72.11	16.27	72.53	14.22	- 0.42	11.47	- 0.28	N.S.

Table 4.23. Paired T-Tests of Change Scores on Objective Tests for All "Tutorial" Sections (12, 04, 09, 10, 21, 03, 15, and 16) in Derivative Comparison 3 (see Table 2.2).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	0.03	N.S.
Sentence Patterns	0.14	N.S.
Paragraph Patterns	0.00	N.S.
Total Score	0.16	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	2.41	N.S.
Daly-Miller Writing Apprehension Test	0.19	N.S.

Table 4.24. Comparison of Posttest Mean Scores on Objective Test Data with Pretest Scores Covaried for "Conventional" and "Tutorial" Sections in Derivative Comparison 3 (see Table 2.2).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	439.63	141.05	421.13	127.88	18.50	120.14	1.73	N.S.
Words/T-units	15.18	3.08	15.23	2.86	- 0.05	3.54	-0.16	N.S.
Words/Clauses	8.80	1.37	9.00	1.56	- 0.20	1.93	-1.17	N.S.
Ratio of TUs with FFM	0.052	0.06	0.054	0.07	- 0.002	0.08	-0.31	N.S.
Ratio of Words in FFM	0.030	0.04	0.031	0.04	- 0.001	0.05	-0.17	N.S.
Holistic Score	4.69	1.54	5.17	1.64	- 0.48	1.94	-2.76	.007
A_2B_2 Topics								
Words	438.44	136.40	456.74	128.54	-18.29	115.27	-1.78	N.S.
Words/T-units	16.95	3.62	16.66	3.42	0.29	4.41	0.74	N.S.
Words/Clauses	10.55	1.63	10.23	1.72	0.31	2.15	1.64	N.S.
Ratio of TUs with FFM	0.041	0.05	0.039	0.05	0.002	0.07	0.38	N.S.
Ratio of Words with FFM	0.021	0.03	0.0205	0.03	0.0005	0.04	0.13	N.S.
Holistic Score	4.61	1.82	5.31	1.89	- 0.70	2.29	-3.43	.001

Table 4.25. Paired T-Tests (N=126) of Change Scores on Writing Samples for All "Analytic" Sections (01, 02, 05, 06, 09, 10, 13, 14, 15, 16, 17, 18, 19, and 20) in Derivative Comparison 4 (see Table 2.2).

Factor	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	405.35	164.70	450.61	153.20	-45.26	98.73	- 3.37	.001
Words/T-units	15.46	2.90	15.41	2.57	0.05	3.29	0.11	N.S.
Words/Clauses	9.20	1.76	9.10	1.25	0.10	1.60	0.46	N.S.
Ratio of TUs with FFM	0.037	0.05	0.038	0.04	- 0.001	0.05	- 0.20	N.S.
Ratio of Words in FFM	0.020	0.03	0.023	0.03	- 0.003	0.04	- 0.63	N.S.
Holistic Score	4.13	1.63	4.67	1.81	- 0.54	2.22	- 1.78	N.S.
A_2B_2 Topics								
Words	438.54	132.91	525.61	167.03	-87.07	156.05	- 4.10	.001
Words/T-units	16.99	3.61	17.26	3.27	- 0.27	3.95	- 0.51	N.S.
Words/Clauses	10.52	1.72	10.78	2.08	- 0.26	2.13	- 0.89	N.S.
Ratio of TUs with FFM	0.035	0.04	0.070	0.06	- 0.035	0.08	- 3.47	.001
Ratio of Words in FFM	0.018	0.02	0.041	0.04	- 0.023	0.04	- 4.04	.001
Holistic Score	4.13	1.74	4.91	2.05	- 0.78	2.50	- 2.29	.026

Table 4.26. Paired T-Tests (N=54) of Change Scores on Writing Samples for All "Synthetic" Sections (12, 04, 17, 18, 21, and 03) in Derivative Comparison 4 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	9.70	.001
T-Unit Length	0.06	N.S.
Clause Length	0.01	N.S.
Ratio of T-Units with Final Free Modifiers	1.72	N.S.
Ratio of Words in Final Free Modifiers	1.00	N.S.
Holistic Score	1.76	N.S.

Table 4.27. Comparison of Posttest Mean Scores on A_1B_1 Topic Data with Pretest Scores Covaried for "Analytic" and "Synthetic" Sections in Derivative Comparison 4 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	13.17	.001
T-Unit Length	1.22	N.S.
Clause Length	3.66	N.S.
Ratio of T-Units with Final Free Modifiers	12.85	.001
Ratio of Words in Final Free Modifiers	14.63	.001
Holistic Score	0.89	N.S.

Table 4.28 Comparison of Posttest Mean Scores on A_2B_2 Topic Data with Pretest Scores Covaried for "Analytic" and "Synthetic" Sections in Derivative Comparison 4 (see Table 2.2).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	116	18.93	3.24	22.16	3.99	- 3.23	4.09	-8.50	.001
Sentence Patterns	116	16.57	3.09	16.96	2.94	- 0.39	3.32	-1.26	N.S.
Paragraph Patterns	116	11.62	2.56	11.41	1.77	0.21	3.02	0.74	N.S.
Total Writing	116	47.12	5.89	50.53	6.57	- 3.41	6.48	-5.67	.001
McGraw-Hill Reading Test									
Reading Comprehension	112	20.24	3.25	21.13	3.29	- 0.88	3.87	-2.41	.017
Daly-Miller Writing Apprehension Test	114	69.62	16.07	69.23	16.18	- 0.62	11.01	-0.60	N.S.

Table 4.29. Paired T-Tests of Change Scores on Objective Tests for All "Analytic" Sections (01, 02, 05, 06, 09, 10, 13, 14, 15, 16, 17, 18, 19, and 20) in Derivative Comparison 4 (see Table 2.2).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	49	18.69	3.66	21.65	4.11	- 2.96	3.35	-6.19	.001
Sentence Patterns	49	15.39	3.30	17.22	4.09	- 1.84	3.15	-4.08	.001
Paragraph Patterns	49	11.67	2.10	11.33	1.75	0.35	2.11	1.15	N.S.
Total Writing	49	45.76	7.34	50.20	8.34	- 4.45	5.73	-5.43	.001
McGraw-Hill Reading Test									
Reading Comprehension	47	19.79	5.39	20.34	4.13	- 0.55	5.09	-0.74	N.S.
Daly-Miller Writing Apprehension Test	45	70.68	16.29	70.60	16.43	0.08	11.82	0.01	N.S.

Table 4.30. Paired T-Tests of Change Scores on Objective Tests for All "Synthetic" Sections (12, 04, 07, 08, 21, and 03) in Derivative Comparison 4 (see Table 2.2).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	0.39	N.S.
Sentence Patterns	3.06	N.S.
Paragraph Patterns	0.10	N.S.
Total Score	0.30	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	1.22	N.S.
Daly-Miller Writing Apprehension Test	0.04	N.S.

Table 4.31. Comparison of Posttest Mean Scores on Objective Test Data with Pretest Scores Covaried for "Analytic" and "Synthetic" Sections in Derivative Comparison 4 (see Table 2.2).

Factor	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	435.14	127.86	419.89	126.73	15.25	115.33	0.79	N.S.
Words/T-units	14.60	3.12	15.07	2.86	- 0.47	3.55	-0.79	N.S.
Words/Clauses	8.46	1.07	8.75	1.30	- 0.29	1.29	-1.36	N.S.
Ratio of TUs with FFM	0.057	0.06	0.052	0.07	0.005	0.08	0.39	N.S.
Ratio of Words in FFM	0.033	0.04	0.027	0.04	0.006	0.05	0.69	N.S.
Holistic Score	5.08	1.54	5.11	1.72	- 0.03	1.86	-0.09	N.S.
A_2B_2 Topics								
Words	427.33	120.48	462.69	124.79	-35.36	110.93	-1.91	N.S.
Words/T-units	16.70	3.17	16.27	2.83	0.43	3.64	0.71	N.S.
Words/Clauses	10.38	1.35	10.18	1.36	0.02	1.42	0.82	N.S.
Ratio of TUs with FFM	0.038	0.05	0.041	0.05	- 0.003	0.07	-0.22	N.S.
Ratio of Words in FFM	0.018	0.03	0.022	0.03	- 0.004	0.04	-0.56	N.S.
Holistic Score	4.31	1.67	5.17	1.75	- 0.86	2.24	-2.30	.027

Table 4.32. Paired T-Tests (N=36) of Change Scores on Writing Samples for All "Analytic-Tutorial" Sections (09, 10, 15, and 16) in Derivative Comparison 5 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	0.01	N.S.
T-Unit Length	0.00	N.S.
Clause Length	0.84	N.S.
Ratio of T-Units with Final Free Modifiers	0.18	N.S.
Ratio of Words in Final Free Modifiers	0.55	N.S.
Holistic Score	0.55	N.S.

Table 4.33. Comparison of Posttest Mean Scores on A_1B_1 Topic Data with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections in Derivative Comparison 5 (see Table 2.2).

Factor	F-ratio	Level of Significance
Essay Length	0.77	N.S.
T-Unit Length	0.52	N.S.
Clause Length	0.01	N.S.
Ratio of T-Units with Final Free Modifiers	0.10	N.S.
Ratio of Words in Final Free Modifiers	0.14	N.S.
Holistic Score	0.07	N.S.

Table 4.34. Comparison of Posttest Mean Scores on A_2B_2 Topics Data with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections in Derivative Comparison 5 (see Table 2.2).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	36	18.97	2.89	23.00	3.11	- 4.03	3.55	-6.81	.001
Sentence Patterns	36	16.39	3.37	17.08	3.25	- 0.69	3.74	-1.11	N.S.
Paragraph Patterns	36	12.00	1.76	11.39	1.46	0.61	2.02	1.82	N.S.
Total Writing	36	47.36	5.89	51.47	5.32	- 4.11	5.41	- 4.56	.001
McGraw-Hill Reading Test									
Reading Comprehension	30	20.13	3.33	20.20	4.18	- 0.07	4.56	-0.08	N.S.
Daly-Miller Writing Apprehension Test	30	73.07	15.38	73.23	13.04	- 0.17	9.94	-0.09	N.S.

Table 4.35. Paired T-Tests of Change Scores of Objective Tests for All "Analytic-Tutorial" Sections (09. 10, 15, and 16) in Derivative Comparison 5 (see Table 2.2).

1

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	2.56	N.S.
Sentence Patterns	0.27	N.S.
Paragraph Patterns	0.03	N.S.
Total Score	1.01	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	3.37	N.S.
Daly-Miller Writing Apprehension Test	0.04	N.S.

Table 4.36. Comparison of Posttest Mean Scores on Objective Test Data with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections in Derivative Comparison 5 (see Table 2.2).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	521.22	163.42	457.17	160.47	64.06	146.06	1.86	N.S.
Words/T-units	16.02	3.58	16.35	3.26	- 0.33	4.22	-0.33	N.S.
Words/Clauses	9.33	1.09	9.31	1.81	0.02	2.53	0.04	N.S.
Ratio of TUs with FFM	0.066	0.05	0.093	0.09	- 0.016	0.10	-1.07	N.S.
Ratio of Words in FFM	0.041	0.04	0.057	0.06	- 0.016	0.07	-0.91	N.S.
Holistic Score	4.94	1.73	5.94	1.80	- 1.00	2.17	-1.96	N.S.
A_2B_2 Topics								
Words	508.83	124.57	531.00	120.12	-22.17	111.43	-0.84	N.S.
Words/T-units	16.32	3.27	17.95	3.94	- 1.63	3.88	-1.78	N.S.
Words/Clauses	10.62	1.67	10.92	1.50	- 0.30	2.18	-0.59	N.S.
Ratio of TUs with FFM	0.063	0.05	0.053	0.07	0.010	0.09	0.48	N.S.
Ratio of Words in FFM	0.037	0.03	0.027	0.04	0.010	0.05	0.82	N.S.
Holistic Score	5.39	2.09	6.39	1.58	- 1.00	2.52	-1.68	N.S.

Table 5.1. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Conventional" Sections (01, 02) in Principal Comparison 1 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A$_1$B$_1$ Topics								
Words	444.06	128.25	479.56	154.48	-35.50	112.90	-1.33	N.S.
Words/T-units	15.69	3.54	16.00	2.53	- 0.30	3.61	-0.36	N.S.
Words/Clauses	9.77	2.07	9.24	1.08	0.53	1.99	1.12	N.S.
Ratio of TUs with FFM	0.052	0.05	0.038	0.04	0.014	0.06	1.04	N.S.
Ratio of Words in FFM	0.031	0.04	0.024	0.03	0.007	0.04	0.67	N.S.
Holistic Score	4.39	1.79	4.44	1.85	- 0.06	1.95	-0.12	N.S.
A$_2$B$_2$ Topics								
Words	473.50	125.83	471.11	134.90	2.39	146.63	0.07	N.S.
Words/T-units	17.20	4.18	16.34	2.82	0.86	2.93	1.24	N.S.
Words/Clauses	10.27	1.75	10.46	1.71	- 0.19	1.58	-0.52	N.S.
Ratio of TUs with FFM	0.029	0.03	0.075	0.06	- 0.046	0.06	-3.61	.002
Ratio of Words in FFM	0.018	0.02	0.054	0.05	- 0.036	0.04	-3.56	.002
Holistic Score	4.22	1.48	5.33	2.30	- 1.11	2.42	-1.95	N.S.

Table 5.2. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Synthetic-Laboratory-Tutorial" Sections (04, 12) in Principal Comparison 1 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	3.07	N.S.
T-Unit Length	0.08	N.S.
Clause Length	0.01	N.S.
Ratio of T-Units with Final Free Modifiers	4.96	.033
Ratio of Words in Final Free Modifiers	4.56	.040
Holistic Score	4.91	.034

Table 5.3. Comparison of Posttest Mean Scores on A_1B_1 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Laboratory-Tutorial" Sections (01, 02 vs. 12,04) in Principal Comparison 1 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	10.58	.003
T-Unit Length	4.53	.045
Clause Length	0.46	N.S.
Ratio of T-Units with Final Free Modifiers	1.51	N.S.
Ratio of Words in Final Free Modifiers	3.85	N.S.
Holistic Score	1.56	N.S.

Table 5.4. Comparison of Posttest Mean Scores on A_2B_2 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Laboratory-Tutorial" Sections (01, 02, vs. 12, 04) In Principal Comparison 1 (see Table 2.1).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	15	19.53	2.64	21.80	4.11	- 2.27	3.97	-2.21	.044
Sentence Patterns	15	16.07	2.87	16.20	2.37	- 0.13	3.09	-0.17	N.S.
Paragraph Patterns	15	10.87	3.70	11.27	1.98	- 0.40	4.37	-0.35	N.S.
Total Writing	15	46.47	5.14	49.27	6.34	- 2.80	6.49	-1.67	N.S.
McGraw-Hill Reading Test									
Reading Comprehension	15	21.07	2.55	21.20	3.17	- 0.13	4.21	-0.12	N.S.
Daly-Miller Writing Apprehension Test	16	63.63	11.04	66.19	18.29	- 2.56	13.37	-0.77	N.S.

Table 5.5. Paired T-Tests of Change Scores on Objective Tests for "Analytic-Conventional" Sections (01, 02) in Principal Comparison 1 (see Table 2.1).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	17	19.65	3.69	21.53	4.03	- 1.88	3.14	-2.47	.025
Sentence Patterns	17	15.76	3.73	17.59	4.62	- 1.82	2.86	-2.63	.018
Paragraph Patterns	17	11.18	2.51	11.35	2.21	- 0.17	2.40	-0.30	N.S.
Total Writing	17	46.59	8.23	50.47	10.05	- 3.88	5.51	-2.91	.010
McGraw-Hill Reading Test									
Reading Comprehension	16	18.88	6.86	19.81	3.41	- 0.94	6.57	-0.58	N.S.
Daly-Miller Writing Apprehension Test	14	65.71	18.88	67.00	14.41	- 1.29	10.63	-0.45	N.S.

Table 5.6. Paired T-Tests (N=18) of Change Scores on Objective Tests for "Synthetic-Laboratory-Tutorial" Sections (04, 12) in Principal Comparison 1 (see Table 2.1).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	0.08	N.S.
Sentence Patterns	2.49	N.S.
Paragraph Patterns	0.05	N.S.
Total Score	0.26	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	0.80	N.S.
Daly-Miller Writing Apprehension Test	0.03	N.S.

Table 5.7. Comparison of Posttest Mean Scores on Objective Tests with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Laboratory-Tutorial" Sections (01, 02 vs. 12, 04) in Principal Comparison 1 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	500.61	128.00	483.72	114.26	16.89	137.96	0.52	N.S.
Words/T-units	15.69	3.09	15.44	2.48	0.25	3.98	0.26	N.S.
Words/Clauses	9.16	1.60	9.38	1.79	- 0.22	2.24	-0.42	N.S.
Ratio of TUs with FFM	0.043	0.05	0.039	0.06	0.004	0.05	0.40	N.S.
Ratio of Words in FFM	0.028	0.03	0.022	0.04	0.006	0.03	0.85	N.S.
Holistic Score	4.83	1.20	5.17	1.65	- 0.33	2.14	-0.66	N.S.
A_2B_2 Topics								
Words	532.94	139.42	494.94	126.72	38.00	115.08	1.40	N.S.
Words/T-units	18.47	5.70	17.88	3.37	0.59	6.21	0.40	N.S.
Words/Clauses	10.81	2.32	10.82	1.95	- 0.01	2.46	-0.02	N.S.
Ratio of TUs with FFM	0.032	0.04	0.045	0.06	- 0.013	0.08	-0.69	M.S.
Ratio of Words in FFM	0.013	0.02	0.023	0.04	- 0.010	0.04	-1.07	N.S.
Holistic Score	4.72	1.78	5.17	2.07	- 0.44	2.04	-0.93	N.S.

Table 5.8. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Conventional" Sections (05, 06) in Principal Comparison 2 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	465.00	160.29	480.39	152.42	- 15.39	82.95	-0.79	N.S.
Words/T-units	14.43	2.80	15.54	3.00	- 1.11	3.13	-1.50	N.S.
Words/Clauses	8.70	1.53	9.18	1.25	- 0.49	1.19	-1.75	[.099]
Ratio of TUs with FFM	0.020	0.02	0.031	0.03	- 0.012	0.04	-1.21	N.S.
Ratio of Words in FFM	0.011	0.02	0.020	0.02	- 0.009	0.03	-1.29	N.S.
Holistic Score	4.44	1.29	5.11	1.91	- 0.67	2.50	-1.13	N.S.
A_2B_2 Topics								
Words	458.94	152.18	575.94	189.72	-117.00	125.87	-3.94	.001
Words/T-units	15.86	3.20	17.50	3.70	- 1.64	3.51	-1.98	[.065]
Words/Clauses	10.57	1.89	10.74	1.38	- 0.17	1.37	-0.53	N.S.
Ratio of TUs with FFM	0.032	0.04	0.064	0.05	- 0.032	0.06	-2.38	.029
Ratio of Words in FFM	0.018	0.03	0.037	0.03	- 0.019	0.04	-2.22	.041
Holistic Score	4.33	1.72	5.06	1.83	- 0.72	2.42	-1.27	N.S.

Table 5.9. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Synthetic-Conventional" Sections (07, 08) in Prinicpal Comparison 2 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	0.31	N.S.
T-Unit Length	0.14	N.S.
Clause Length	0.01	N.S.
Ratio of T-Units with Final Free Modifiers	0.32	N.S.
Ratio of Words in Final Free Modifiers	0.76	N.S.
Holistic Score	0.05	N.S.

Table 5.10. Comparison of Posttest Mean Scores on A_1B_1 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections (05, 06 vs. 07, 08) in Principal Comparison 2 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	11.60	.002
T-Unit Length	0.01	N.S.
Clause Length	0.00	N.S.
Ratio of T-Units with Final Free Modifiers	1.05	N.S.
Ratio of Words in Final Free Modifiers	1.44	N.S.
Holistic Score	0.00	N.S.

Table 5.11. Comparison of Posttest Mean Scores on A_2B_2 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections (05, 06, vs. 07, 08) in Principal Comparison 2 (see Table 2.1).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	16	18.63	3.90	22.00	2.53	- 3.38	3.38	-3.99	.001
Sentence Patterns	16	16.25	2.46	18.13	2.34	- 1.88	2.36	-3.17	.006
Paragraph Patterns	16	11.50	3.46	11.56	1.50	- 0.06	3.97	-0.06	N.S.
Total Writing	16	46.38	5.58	57.69	4.19	- 5.31	4.84	-4.39	.001
McGraw-Hill Reading Test									
Reading Comprehension	17	18.94	3.58	20.71	2.66	- 1.76	3.80	-1.91	N.S.
Daly-Miller Writing Apprehension Test	18	64.72	14.74	66.06	17.49	- 1.33	8.56	-0.66	N.S.

Table 5.12. Paired T-Tests of Change Scores on Objective Tests for "Analytic-Conventional" Sections (05, 06) in Principal Comparison 2 (see Table 2.1).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	16	19.37	2.31	23.32	2.44	- 3.94	3.24	-4.87	.001
Sentence Patterns	16	15.37	1.93	17.94	2.54	- 2.56	2.96	-3.46	.004
Paragraph Patterns	16	12.50	1.27	11.19	1.11	1.31	1.92	2.73	.015
Total Writing	16	47.25	4.12	52.44	4.84	- 5.19	6.18	-3.36	.004
McGraw-Hill Reading Test									
Reading Comprehension	17	20.94	4.15	20.53	3.48	0.41	4.54	0.37	N.S.
Daly-Miller Writing Apprehension Test	18	70.17	14.89	69.11	17.70	1.06	9.76	0.46	N.S.

Table 5.13. Paired T-Tests of Change Scores on Objective Tests for "Synthetic-Conventional" Sections (07, 08) in Principal Comparison 2 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	1.73	N.S.
T-Unit Length	0.03	N.S.
Clause Length	0.36	N.S.
Ratio of T-Units with Final Free Modifiers	0.10	N.S.
Ratio of Words in Final Free Modifiers	0.37	N.S.
Holistic Score	0.61	N.S.

Table 5.14. Comparison of Posttest Mean Scores on Objective Tests with Pretest Scores Covaried for "Analytic-Conventional" and "Synthetic-Conventional" Sections (05, 06 vs. 07, 08) in Principal Comparison 2 (see Table 2.1).

Factor	Pretest		Posttest					
	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	371.33	110.84	384.67	117.54	-13.33	94.75	-0.60	N.S.
Words/T-units	15.21	3.28	13.64	2.75	1.57	3.03	2.20	.042
Words/Clauses	8.12	1.12	8.08	1.53	0.44	1.31	0.14	N.S.
Ratio of TUs with FFM	0.061	0.07	0.030	0.04	0.031	0.07	1.93	N.S.
Ratio of Words in FFM	0.034	0.04	0.015	0.03	0.018	0.05	1.66	N.S.
Holistic Score	4.06	1.66	5.50	1.79	- 1.44	2.15	-2.85	.011
A_2B_2 Topics								
Words	385.11	135.38	393.17	142.23	- 8.06	122.48	-0.28	N.S.
Words/T-units	16.86	3.62	14.08	3.14	2.77	4.79	2.45	.025
Words/Clauses	10.60	1.82	8.59	1.49	2.01	2.64	3.22	.005
Ratio of TUs with FFM	0.037	0.04	0.035	0.05	0.002	0.05	0.19	N.S.
Ratio of Words in FFM	0.022	0.03	0.019	0.03	0.003	0.04	0.37	N.S.
Holistic Score	4.89	2.05	5.28	1.90	- 0.39	2.59	-0.64	N.S.

Table 5.15. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Conventional" Sections (17, 18) in Principal Comparison 3 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	407.83	122.05	395.50	91.34	12.33	103.31	0.51	N.S.
Words/T-units	16.26	2.65	15.64	2.43	0.62	2.76	0.96	N.S.
Words/Clauses	9.15	1.75	9.37	1.59	- 0.22	2.65	-0.35	N.S.
Ratio of TUs with FFM	0.051	0.06	0.067	0.07	- 0.016	0.09	-0.73	N.S.
Ratio of Words in FFM	0.027	0.04	0.039	0.05	- 0.012	0.06	-0.80	N.S.
Holistic Score	4.33	1.14	4.72	1.02	- 0.39	1.20	-1.38	N.S.
A_2B_2 Topics								
Words	405.83	126.90	462.67	114.04	-56.83	125.48	-1.92	N.S.
Words/T-units	17.23	2.55	16.82	2.65	0.41	3.22	0.54	N.S.
Words/Clauses	10.55	1.43	10.04	1.59	0.51	1.93	1.11	N.S.
Ratio of TUs with FFM	0.036	0.05	0.030	0.04	0.006	0.05	0.54	N.S.
Ratio of Words in FFM	0.017	0.03	0.016	0.02	0.001	0.03	0.13	N.S.
Holistic Score	4.83	1.92	5.44	2.04	- 0.61	2.66	-0.97	N.S.

Table 5.16. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Syllabus-Conventional" Sections (19, 20) in Principal Comparison 3 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	0.10	N.S.
T-Unit Length	3.97	N.S.
Clause Length	4.93	.033
Ratio of T-Units with Final Free Modifiers	3.95	N.S.
Ratio of Words in Final Free Modifiers	3.05	N.S.
Holistic Score	3.22	N.S.

Table 5.17. Comparison of Posttest Mean Scores on A_1B_1 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Syllabus-Conventional" Sections (17, 18 vs. 19, 20) in Principal Comparison 3 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	2.53	N.S.
T-Unit Length	7.63	.009
Clause Length	7.69	.009
Ratio of T-Units with Final Free Modifiers	0.11	N.S.
Ratio of Words in Final Free Modifiers	0.02	N.S.
Holistic Score	0.07	N.S.

Table 5.18. Comparison of Posttest Mean Scores on A_2B_2 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Syllabus-Conventional" Sections (17, 18 vs. 19, 20) in Principal Comparison 3 (see Table 2.1).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	17	19.35	4.17	21.06	5.96	- 1.71	5.22	-1.35	N.S.
Sentence Patterns	17	17.41	2.69	17.12	2.80	0.29	3.09	0.39	N.S.
Paragraph Patterns	17	11.47	3.06	11.47	1.94	0.00	3.48	0.00	N.S.
Total Writing	17	48.24	6.70	49.65	8.17	- 1.41	8.34	-0.70	M.S.
McGraw-Hill Reading Test									
Reading Comprehension	15	19.47	3.50	21.73	2.96	- 2.27	4.18	-2.10	N.S.
Daly-Miller Writing Apprehension Test	17	70.53	18.34	70.35	17.14	0.18	15.94	0.05	N.S.

Table 5.19. Paired T-Tests of Change Scores on Objective Tests for "Analytic-Conventional" Sections (17, 18) in Principal Comparison 3 (see Table 2.1).

Factor	N	Pretest		Posttest		Difference	S.D.	T-Value	Significance
		Mean	S.D.	Mean	S.D				
McGraw-Hill Writing Test									
Language Mechanics	17	18.35	3.10	22.65	3.59	- 4.29	4.15	-4.27	.001
Sentence Patterns	17	16.53	4.13	16.65	3.10	- 0.12	3.24	-0.15	N.S.
Paragraph Patterns	17	11.41	1.66	11.59	1.73	- 0.17	2.60	-0.28	N.S.
Total Writing	17	46.29	7.04	50.88	6.60	- 4.59	7.00	-2.70	.016
McGraw-Hill Reading Test									
Reading Comprehension	18	20.56	3.26	21.67	3.25	- 1.11	2.76	-1.71	N.S.
Daly-Miller Writing Apprehension Test	18	70.56	18.69	71.72	16.75	- 1.17	10.37	-0.48	N.S.

Table 5.20. Paired T-Tests of Change Scores for Objective Tests for "Analytic-Syllabus-Conventional" Sections (19, 20) in Prinicpal Comparison 3 (see Table 2.1).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	1.90	N.S.
Sentence Patterns	0.01	N.S.
Paragraph Patterns	0.03	N.S.
Total Score	0.80	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	0.24	N.S.
Daly-Miller Writing Apprehension Test	0.11	N.S.

Table 5.21. Comparison of Posttest Mean Scores on Objective Tests with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Syllabus-Conventional" Sections (17, 18 vs. 19, 20) in Principal Comparison 3 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A$_1$B$_1$ Topics								
Words	406.17	155.30	387.11	131.74	19.06	124.08	0.65	N.S.
Words/T-units	13.87	2.14	15.40	2.95	- 1.53	3.06	-2.12	.049
Words/Clauses	8.93	1.45	9.38	1.23	- 0.45	1.93	-0.98	N.S.
Ratio of TUs with FFM	0.032	0.04	0.050	0.04	- 0.018	0.05	-1.57	N.S.
Ratio of Words in FFM	0.018	0.03	0.031	0.03	- 0.013	0.04	-1.50	N.S.
Holistic Score	4.50	1.76	4.61	1.46	- 0.11	1.81	-0.26	N.S.
A$_2$B$_2$ Topics								
Words	381.72	122.56	390.00	95.99	- 8.28	100.86	-0.35	N.S.
Words/T-units	16.38	3.07	17.38	3.81	- 1.00	4.01	-1.05	N.S.
Words/Clauses	10.51	1.47	10.89	1.71	- 0.38	2.02	-0.80	N.S.
Ratio of TUs with FFM	0.043	0.07	0.029	0.03	0.014	0.06	0.96	N.S.
Ratio of Words in FFM	0.023	0.03	0.016	0.02	0.007	0.03	0.85	N.S.
Holistic Score	3.83	1.20	4.55	1.89	- 0.72	1.87	-1.64	N.S.

Table 5.22. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Conventional" Sections (13, 14) in Principal Comparison 4 (see Table 2.1).

	Pretest		Posttest					
Factor	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	469.50	129.52	431.89	125.77	37.61	100.98	1.58	N.S.
Words/T-units	15.13	3.65	15.72	3.05	- 0.59	3.53	-0.71	N.S.
Words/Clauses	8.57	0.86	9.23	1.59	- 0.66	1.28	-2.19	N.S.
Ratio of TUs with FFM	0.082	0.06	0.061	0.08	0.210	0.10	0.91	N.S.
Ratio of Words in FFM	0.050	0.04	0.032	0.04	0.018	0.06	1.36	N.S.
Holistic Score	5.56	1.38	5.17	1.72	0.39	1.88	0.88	N.S.
A_2B_2 Topics								
Words	447.22	105.00	468.89	109.57	-21.67	105.62	-0.87	N.S.
Words/T-units	17.79	3.24	16.60	2.69	1.19	3.77	1.34	N.S.
Words/Clauses	10.61	1.40	10.23	1.29	0.38	1.37	1.17	N.S.
Ratio of TUs with FFM	0.045	0.06	0.067	0.05	- 0.022	0.07	-1.29	N.S.
Ratio of Words in FFM	0.019	0.03	0.036	0.04	- 0.017	0.04	-1.69	N.S.
Holistic Score	4.17	1.72	4.89	1.71	- 0.72	2.42	-1.27	N.S.

Table 5.23. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Tutorial" Sections (15, 16) in Principal Comparison 4 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	0.46	N.S.
T-Unit Length	0.03	N.S.
Clause Length	0.01	N.S.
Ratio of T-Units with Final Free Modifiers	0.06	N.S.
Ratio of Words in Final Free Modifiers	0.05	N.S.
Holistic Score	0.15	N.S.

Table 5.24. Comparison of Posttest Mean Scores on A_1B_1 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections (13, 14, vs. 15, 16) in Principal Comparison 4 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	2.33	N.S.
T-Unit Length	1.14	N.S.
Clause Length	2.02	N.S.
Ratio of T-Units with Final Free Modifiers	7.82	.009
Ratio of Words in Final Free Modifiers	4.36	.045
Holistic Score	0.21	N.S.

Table 5.25. Comparison of Posttest Mean Scores on A_2B_2 Writing Samples with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections (13, 14, vs. 15, 16) in Principal Comparison 4 (see Table 2.1).

		Pretest		Posttest					
Factor	N	Mean	S.D.	Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	15	18.73	3.17	21.40	4.79	- 2.67	4.48	-2.30	.037
Sentence Patterns	15	16.93	2.43	16.33	3.22	0.60	4.48	0.66	N.S.
Paragraph Patterns	15	12.00	2.14	11.20	2.48	0.80	2.31	1.34	N.S.
Total Writing	15	47.67	5.16	48.94	9.48	- 1.27	7.20	-0.68	N.S.
McGraw-Hill Reading Test									
Reading Comprehension	17	21.35	2.89	22.00	2.26	- 0.64	2.89	-0.92	N.S.
Daly-Miller Writing Apprehension Test	15	72.87	14.14	71.73	17.05	1.13	7.94	0.55	N.S.

Table 5.26. Paired T-Tests of Change Scores on Objective Tests for "Analytic-Conventional" Sections (13, 14) in Principal Comparison 4 (see Table 2.1).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	18	19.56	3.00	23.50	3.00	- 3.94	3.33	-5.02	.001
Sentence Patterns	18	15.89	3.25	16.83	3.12	- 0.94	3.28	-1.22	N.S.
Paragraph Patterns	18	12.17	1.50	11.56	1.25	0.61	1.50	1.73	N.S.
Total Writing	18	47.61	5.56	51.89	5.29	- 4.28	4.13	-4.40	.001
McGraw-Hill Reading Test									
Reading Comprehension	16	20.63	3.26	21.13	2.92	- 0.50	3.33	-0.60	N.S.
Daly-Miller Writing Apprehension Test	16	70.25	15.71	71.63	15.03	- 1.38	8.46	-0.65	N.S.

Table 5.27. Paired T- Tests of Change Scores on Objective Tests for "Analytic-Tutorial" Sections (15, 16) in Principal Comparison 4 (see Table 2.1).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	1.71	N.S.
Sentence Patterns	0.81	N.S.
Paragraph Patterns	0.21	N.S.
Total Score	2.21	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	0.55	N.S.
Daly-Miller Writing Apprehension Test	0.60	N.S.

Table 5.28. Comparison of Posttest Mean Scores on Objective Tests with Pretest Scores Covaried for "Analytic-Conventional" and "Analytic-Tutorial" Sections (13, 14 vs. 15, 16) in Principal Comparison 4 (see Table 2.1).

Factor	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	400.78	119.93	407.89	130.17	- 7.11	127.00	-0.24	N.S.
Words/T-units	14.06	2.49	14.41	2.58	- 0.35	3.66	-0.40	N.S.
Words/Clauses	8.36	1.27	8.28	0.70	0.08	1.23	0.26	N.S.
Ratio of TUs with FFM	0.032	0.05	0.042	0.05	- 0.010	0.06	-0.76	N.S.
Ratio of Words in FFM	0.015	0.03	0.023	0.03	- 0.008	0.03	-1.26	N.S.
Holistic Score	4.61	1.58	5.06	1.77	- 0.44	1.79	-1.05	N.S.
A_2B_2 Topics								
Words	407.44	134.24	456.50	141.33	-49.06	117.40	-1.77	N.S.
Words/T-units	15.61	2.78	15.94	3.01	- 0.33	3.44	-0.41	N.S.
Words/Clauses	10.14	1.30	10.13	1.46	0.01	1.49	0.04	N.S.
Ratio of TUs with FFM	0.032	0.05	0.015	0.03	0.017	0.05	1.35	N.S.
Ratio of Words in FFM	0.018	0.03	0.008	0.02	0.010	0.03	1.22	N.S.
Holistic score	4.44	1.65	5.44	1.79	- 1.00	2.11	-2.01	N.S.

Table 5.29. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Analytic-Tutorial" Sections (09, 19) in Principal Comparison 5 (see Table 2.1).

Factor	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D	Difference	S.D.	T-Value	Significance
A_1B_1 Topics								
Words	307.00	163.75	391.89	143.46	- 84.89	89.74	-4.01	.001
Words/T-units	16.25	2.00	14.69	2.08	1.56	2.65	2.50	.023
Words/Clauses	9.15	1.56	8.89	1.44	0.26	1.43	0.77	N.S.
Ratio of TUs with FFM	0.038	0.06	0.045	0.05	- 0.007	0.06	-0.49	N.S.
Ratio of Words in FFM	0.018	0.03	0.026	0.03	- 0.008	0.04	-0.86	N.S.
Holistic Score	3.56	1.69	4.44	1.69	- 0.89	2.22	-1.70	N.S.
A_2B_2 Topics								
Words	383.17	105.40	529.78	164.15	-146.61	159.38	-3.90	.001
Words/T-units	17.90	3.28	17.95	3.21	- 0.05	4.93	-0.04	N.S.
Words/Clauses	10.73	1.56	11.14	2.90	- 0.41	3.12	-.056	N.S.
Ratio of TUs with FFM	0.043	0.05	0.071	0.08	- 0.012	0.08	-1.12	N.S.
Ratio of Words in FFM	0.019	0.02	0.032	0.04	- 0.012	0.04	-1.24	N.S.
Holistic Score	3.83	2.04	4.33	1.97	- 0.50	2.75	-0.77	N.S.

Table 5.30. Paired T-Tests (N=18) of Change Scores on Writing Samples for "Synthetic-Tutorial" Sections (21, 03) in Principal Comparison 5 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	1.71	N.S.
T-Unit Length	0.05	N.S.
Clause Length	0.78	N.S.
Ratio of T-Units with Final Free Modifiers	0.01	N.S.
Ratio of Words in Final Free Modifiers	0.06	N.S.
Holistic Score	0.25	N.S.

Table 5.31. Comparison of Posttest Mean Scores on A_1B_1 Writing Samples with Pretest Scores Covaried for "Analytic-Tutorial" and "Synthetic-Tutorial" Sections (09, 10 vs. 21, 03) in Principal Comparison 5 (see Table 2.1).

Factor	F-ratio	Level of Significance
Essay Length	3.85	N.S.
T-Unit Length	2.82	N.S.
Clause Length	1.11	N.S.
Ratio of T-Units with Final Free Modifiers	7.75	.009
Ratio of Words in Final Free Modifiers	5.28	.028
Holistic Score	2.57	N.S.

Table 5.32. Comparison of Posttest Mean Scores on A_2B_2 Writing Samples with Pretest Scores Covaried for "Analytic-Tutorial" and "Synthetic-Tutorial" Sections (09, 10 vs. 21, 03) in Principal Comparison 5 (see Table 2.1).

Factor	N	Pretest Mean	Pretest S.D.	Posttest Mean	Posttest S.D.	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	18	18.93	2.75	22.50	3.22	- 4.11	3.85	-4.53	.001
Sentence Patterns	18	16.89	3.50	17.33	3.25	- 0.44	4.23	-0.45	N.S.
Paragraph Patterns	18	11.83	2.01	11.22	1.67	0.61	2.48	1.05	N.S.
Total Writing	18	47.11	6.34	51.06	5.46	- 3.94	6.57	-2.55	.021
McGraw-Hill Reading Test									
Reading Comprehension	14	19.57	3.44	19.14	5.19	0.43	5.77	0.28	N.S.
Daly-Miller Writing Apprehension Test	14	76.29	14.90	75.07	10.57	1.21	11.59	0.39	N.S.

Table 5.33. Paired T-Tests of Change Scores on Objective Tests for "Analytic-Tutorial" Sections (09, 10) in Principal Comparison 5 (see Table 2.1).

Factor	N	Pretest Mean	S.D.	Posttest Mean	S.D	Difference	S.D.	T-Value	Significance
McGraw-Hill Writing Test									
Language Mechanics	16	17.00	4.31	20.13	5.03	- 3.13	3.54	-3.53	.003
Sentence Patterns	16	15.00	3.98	16.13	4.73	- 1.13	3.63	-1.24	N.S.
Paragraph Patterns	16	11.38	2.19	11.44	1.83	- 0.63	1.69	-0.15	N.S.
Total Writing	16	43.38	8.60	47.69	8.90	- 4.31	5.81	-2.97	.010
McGraw-Hill Reading Test									
Reading Comprehension	14	19.43	4.94	20.71	5.61	- 1.29	3.95	-1.21	N.S.
Daly-Miller Writing Apprehension Test	13	76.77	14.27	76.85	15.82	0.08	15.86	0.02	N.S.

Table 5.34. Paired T-Tests of Change Scores on Objective Tests for "Synthetic-Tutorial" Sections (21, 03) in Principal Comparison 5 (see Table 2.1).

Factor	F-ratio	Level of Significance
McGraw-Hill Writing Test		
Language Mechanics	1.43	N.S.
Sentence Patterns	0.03	N.S.
Paragraph Patterns	0.41	N.S.
Total Score	0.26	N.S.
McGraw-Hill Reading Test		
Paragraph Comprehension	0.85	N.S.
Daly-Miller Writing Apprehension Test	0.12	N.S.

Table 5.35. Comparison of Posttest Mean Scores on Objective Tests with Pretest Scores Covaried for "Analytic-Tutorial" and "Synthetic-Tutorial" Sections (09, 10 vs. 21, 03) in Principal Comparison 5 (see Table 2.1).

	Low Apprehensives (N=55) Mean	Low Apprehensives (N=55) S.D.	High Apprehensives (N=55) Mean	High Apprehensives (N=55) S.D.	Difference	F-ratio	Significance
Factor							
Words	479.9	138.7	396.6	139.6	83.3	9.85	.002
Words/T-units	15.9	3.42	14.7	2.78	1.2	4.65	.033
Words/Clauses	9.58	1.59	8.43	1.41	1.55	16.06	.001
Ratio of TUs with FFM	.060	.063	.039	.045	.021	4.32	.040
Ratio of Words in FFM	.038	.042	.023	.030	.015	4.80	.031
Holistic Score	4.92	1.55	4.29	1.56	0.63	4.61	.034

Table 6.1. Comparison of Pretest Mean Scores for A_1B_1 Topic Data for High- and Low-Writing Apprehensives.

	Low Apprehensives (N=55)		High Apprehensives (N=55)				
	Mean	S.D.	Mean	S.D.	Difference	F-ratio	Significance
Factor							
Words	487.4	146.5	413.1	127.6	74.3	8.05	.005
Words/T-units	17.4	4.30	16.6	2.87	0.8	1.52	N.S.
Words/Clauses	10.8	1.85	10.5	1.40	0.5	2.53	N.S.
Ratio of TUs with FFM	.047	.046	.027	.045	.020	5.19	.025
Ratio of Words in FFM	.026	.028	.015	.027	.011	4.47	.037
Holistic Score	4.54	1.90	4.25	1.79	0.29	.683	N.S.

Table 6.2. Comparison of Pretest Mean Scores for A_2B_2 Topic Data for High- and Low-Writing Apprehensives.

Factor	Low Apprehensives (N=55) Mean	S.D.	High Apprehensives (N=55) Mean	S.D.	Difference	F-ratio	Significance
McGraw-Hill Writing Test (Total)	48.4	6.05	46.5	6.03	1.9	2.42	N.S.
SAT Verbal							
Reading Comprehension	48.3	7.18	46.1	7.01	2.2	2.00	N.S.
Vocabulary	48.3	7.99	43.9	8.52	4.4	6.34	.014
Total Verbal	485.6	62.1	451.5	75.9	34.1	6.57	.012
McGraw-Hill Reading Test	20.9	3.47	19.0	4.74	1.9	4.99	.028
English Composition Test	469.1	54.2	430.9	65.8	38.2	11.03	.001
Test of Standard Written English	48.1	6.26	44.8	8.83	3.3	4.11	.046

Table 6.3. Comparison of Pretest Mean Scores Test Data for High- and Low-Writing Apprehensives.

Step	Variable	F-ratio to enter	R^2	R^2 Change
1	Words/clauses 1	21.87	.1171	.1171
2	Total words 1	13.50	.1842	.0671
3	Ratio/words in FFM 2	6.31	.2106	.0304
4	McGraw-Hill Reading	6.25	.2438	.0292
5	Ratio/words in FFM 1	1.34	.2500	.0063
6	Words/T-units 1	1.03	.2548	.0048
7	Ratio/T-units with FFM1	.60	.2576	.0028
8	Words/clauses 2	.54	.2602	.0025
9	English Composition Test	.52	.2625	.0024

The following variables were also entered into the regression equation, but together failed to predict more than 2% of the remaining variance.

10	SAT Verbal
11	Holistic score 1
12	SAT-Reading Comprehension
13	Test of Standard Written English
14	Total words 2
15	McGraw-Hill Writing Test
16	Words/T-units 2
17	Ratio/T-units with FFM 2
18	Holistic score 2

Table 6.4. Stepwise Regression on 18 Variables as Predictors of Writing Apprehension (N=161).

Step	Variable	F to Enter	R^2	R^2 Change
1	Total Words	8.57	.0494	.0494
2	Ratio of Words in FFM	5.41	.0797	.0303

Table 7.1. Stepwise Regression on Five Variables as Predictors of Holistic Score on the A_1B_1 Posttest Essays.

Step	Variable	F to Enter	R^2	R^2 Change
1	Total Words	11.51	.0652	.0652
2	Ratio of Words in FFM	1.29	.0725	.0073
3	Ratio of T-units with FFM	1.15	.0790	.0065

Table 7.2. Stepwise Regression on Five Variables as Predictors of Holistic Scores on the A_2B_2 Posttest Essays.

Step	Variable	F to Enter	R^2	R^2 Change
1	Total Words--A_2B_2	13.81	.0772	.0772
2	Ratio of Words in FFM--A_1B_1	8.44	.1224	.0452

Table 7.3. Stepwise Regression of Ten Variables as Predictors of the Summed Holistic Scores on the A_1B_1 and the A_2B_2 Posttest Essays.

Step	Variable	F to Enter	R^2	R^2 Change
1	Total Words	22.29	.1113	.1113
2	Ratio of Words in FFM	4.87	.1351	.0238
3	Words per T-Unit	2.93	.1493	.0142

Table 7.4. Stepwise Regression on Five Variables as Predictors of the Summed Holistic Scores for the Two Pretest and the Two Posttest Essay Sets.

Step	Variable	F to Enter	R^2	R^2 Change
1	Holistic Score--A_1B_1	10.15	.0580	.0580
2	SAT Verbal	5.97	.0910	.0331
3	Ratio of T-Units with FFM--A_2B_2	2.39	.1041	.0130
4	Ratio of Words in FFM--A_1B_1	1.67	.1133	.0092

Table 7.5. Stepwise Regression on 17 Pretest Variables as Predictors of Course Grade in E 306.

BIBLIOGRAPHY OF WORKS CITED

Adelstein, M.E., and Pival, J.G. The writing commitment. New York: Harcourt Brace Jovanovich, 1976.

Anderson, C.C. The new STEP essay test as a measure of composition ability. Educational and Psychological Measurement, 1960, 20, 95-102.

Barritt, L., and Kroll, B. Some implications of cognitive-developmental psychology for research in composing. In Cooper, C., and Odell, L. (Eds.) Research on composing: some points of departure. Urbana, Ill.: National Council of Teachers of English, 1978.

Becker, A. A tagmemic approach to paragraph analysis. College Composition and Communication, 1965, 16, 237-242.

Blumenthal, J.C. English 3200: a programmed course in grammar and usage. New York: Harcourt Brace Jovanovich, 1972.

Bortz, D.R. The written language patterns of intermediate grade children when writing compositions in three forms: descriptive, expository, and narrative. Unpublished doctoral dissertation, Lehigh University, 1969.

Brereton, J., Perl, S., and Sterling, R. An untitled presentation made during the Annual Meeting of the Conference on College Composition and Communication, 1979.

Britton, J., Burgess, T., Martin, N., McLeod, A., and Rosen, H. The development of writing abilities (11-18). London: Macmillan Education, 1975.

Braddock, R., Lloyd-Jones, R., and Schoer, L. Research in written composition. Champaign, Il.: National Council of Teachers of English, 1963.

Campbell, D.T., and Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Christensen, F. A generative rhetoric of the paragraph. College Composition and Communication, 1965, 16, 144-156.

Christensen, F. *Notes toward a new rhetoric: six essays for teachers*. New York: Harper & Row, 1967.

Christensen, F., and Christensen, B. *Notes toward a new rhetoric*. New York: Harper & Row, 1979.

Combs, W. Further effects of sentence-combining on writing ability. *Research in the Teaching of English*, 1976, *10*, 137-149.

Cooper, C.R., et al. *Writing abilities of regularly admitted freshmen at SUNY/Buffalo*. Buffalo: State University of New York at Buffalo, 1979.

Corder, J.W. *Handbook of current English*. Glenview, Illinois: Scott, Foresman, 1978.

Crowhurst, M. The effect of syntactic complexity on teachers' quality rating of written compositions. Paper presented during the Annual Meeting of the Conference on College Composition and Communication, 1979.

Crowhurst, M. *The effect of audience and mode of discourse on the syntactic complexity of the writing of sixth and tenth graders*. Unpublished doctoral dissertation, University of Minnesota, 1978.

Cummings, D., Herum, J., and Lybert, E. Semantic recurrence and rhetorical form. *Language and Style*, 1971, *4*, 195-207.

Daiker, D., Kerek, A., and Morenberg, M. Sentence combining and syntactic maturity in freshman English. *College Composition and Communication*, 1978, *29*, 36-41.

Daly, J.A. The assessment of social-communicative anxiety via self-reports: a comparison of measures. *Communication Monographs*, 1978, *45*, 204-218. a.

Daly, J.A. The effects of writing apprehension on message encoding. *Journalism Quarterly*, 1977, *54*, 566-572.

Daly, J.A. Writing apprehension and self esteem. Paper presented to the Annual Meeting of the Southwest Educational Research Association, 1980.

Daly, J.A. Writing apprehension and writing competency. Journal of Educational Research, 1978, 72, 10-14. b.

Daly, J.A., and Miller, M.D. The development of a measure of writing apprehension. Research in the Teaching of English, 1975, 9, 242-249. a.

Daly, J.A., and Miller, M.D. Further studies in writing apprehension: SAT scores, success expectations, willingness to take advanced courses and sex differences. Research in the Teaching of English, 1975, 9, 250-256. b.

Daly, J.A., and Shamo, W. Academic decisions as a function of writing apprehension. Research in the Teaching of English, 1978, 12, 119-126.

Daly, J.A., and Shamo, W. Writing apprehension and occupational choice. Journal of Occupational Psychology, 1976, 49, 55-56.

D'Angelo, F.J. A generative rhetoric of the essay. College Composition and Communication, 1974, 25, 388-396.

D'Angelo, F.J. Process and thought in composition. Cambridge, MA: Winthrop, 1980.

Decker, R.E., ed. Patterns of exposition 6. Boston: Little, Brown, 1978.

Diederich, P.E., French, S.W., and Carlton, S.T. Factors in judgments of writing ability. Research Bulletin RB-61-15. Princeton, N.J.: Educational Testing Service, 1961.

Ebel, R.L. Estimation of the reliability of ratings. Psychometrika, 1951, 16, 407-424.

Faigley, L. Another look at sentences. Freshman English News, 1979, 7, 18-21. a.

Faigley, L. Generative rhetoric as a way of increasing syntactic fluency. College Composition and Communication, 1979, 30, 176-181. b.

Faigley, L. The influence of generative rhetoric on the syntactic fluency and writing effectiveness of college freshmen. Research in the Teaching of English, 1979, 13, 197-206. c.

Faigley, L. Problems of analyzing maturity in college and adult writing. In Daiker, D., Kerek, A., and Morenberg, M. (Ed.) Sentence combining and the teaching of writing. Akron, Ohio: University of Akron, 1979. d.

Faigley, L. Names in search of a concept: maturity, fluency, complexity, and growth in written syntax. College Composition and Communication, 1980, 31, in press.

Frederiksen, C. Semantic processing units in understanding texts. In Freedle, R. (Ed.), Discourse production and comprehension. Norwood, N.J.: Ablex, 1977.

Frogner, E. Problems of sentence structure in pupils' themes. English Journal, 1933, 22, 742-749.

Gage, N.L. Handbook of research on teaching. Chicago: Rand McNally, 1963.

Gebhard, A.O. Writing quality and syntax: a transformational analysis of three prose samples. Research in the Teaching of English, 1978, 12, 211-231.

Grady, M. A conceptual rhetoric of the composition. College Composition and Communication, 1971, 22, 348-354.

Hall, D., and Emblen, D. (Eds.). A writer's reader. Boston: Little, Brown, 1976.

Halliday, M.A.K., and Hasan, R. Cohesion in English. London: Longmans, 1976.

Harris, S.L., and Witte, S.P. Sentence combining in a rhetorical framework: directions for further research. In Pringle, I., and Freedman, A. (Eds.), Reinventing the rhetorical tradition, Akron, OH: Canadian Council of Teachers of English and the University of Akron, 1980.

Harris, Z. Discourse analysis: a sample text. Language, 1952, 28, 474-494.

Harvley-Felder, Z.C. Some factors relating to writing apprehension: An exploratory study. Unpublished doctoral dissertation, Univ. of North Carolina, 1978.

Hunt, K.W. Differences in grammatical structures written at three grade levels, the structures to be analyzed by transformational methods. Tallahassee: Florida State University, 1964.

Hunt, K.W. Early blooming and late blooming syntactic structures. In Cooper, C., and Odell, L. (Eds.). Evaluating writing: describing, measuring, judging. Urbana, Ill.: NCTE, 1977.

Hunt, K.W. Grammatical structures written at three grade levels. NCTE Research Report, No. 3. Urbana, Ill.: National Council of Teachers of English, 1965.

Hunt, K.W. Syntactic maturity in schoolchildren and adults. Monographs of the Society for Research in Child Development, No. 134. Chicago: University of Chicago Press, 1970.

Jakobson, R. Linguistics and poetics. In Sebeok, T.A. (Ed.), Style in language. Cambridge, Mass.: M.I.T. Press, 1960.

Jewell, R.M., Cowley, J., Rhum, G., Snider, W., and Schnittjer, C. The effectiveness of college-level instruction in freshman composition. Final report, Project No. 2188. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1969.

Johnson, L.V. Children's writing in three forms of composition. Elementary Education, 1967, 44, 265-269.

Johnson, S.T. Some tentative structures on generative rhetoric. College English, 1969, 31, 155-165.

Karrfalt, D. The generation of paragraphs and larger units. College Composition and Communication, 1968, 19, 211-217.

Kincaid, G.L. Some factors effecting variations in the quality of student's writing. Unpublished doctoral dissertation, Michigan State University, 1953.

Kinneavy, J.L. Sentence combining in a comprehensive language framework. In Daiker, D., Kerek, A., and Morenberg, M. (Eds.) Sentence combining and the teaching of writing. Akron, Ohio: University of Akron, 1979.

Kinneavy, J.L. A theory of discourse. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Kinneavy, J.L., Cope, J.Q., and Campbell, J.W. Aims and audiences in writing. Dubuque, Iowa: Kendall/Hunt, 1976. a.

Kinneavy, J.L., Cope, J.Q. Campbell, J.W. Writing--basic modes of organization. Dubuque, Iowa: Kendall/Hunt, 1976. b.

Kinneavy, J.L. Daly, J.A., Faigley, L., and Witte, S.P. Evaluating the effectiveness of college writing programs: a proposal submitted to the Fund for the Improvement of Postsecondary Education. April, 1980.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W., and Keenan, J.M. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.

Kintsch, W., Kozminsky, E., Streby, W.J., McKoon, G., and Keenan, J.M. Comprehension and recall of text as a function of content variables. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 196-214.

Kintsch, W., and van Dijk, T.A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Koffend, J. The case for alcohol. Atlantic Monthly, Dec., 1979.

LaBrant, L. A study of certain language developments in grades four through twelve, inclusive. Genetic Psychology Monographs, 1933, 14, 387-491.

Larson, R.L. Evaluating instruction in composition: an art not yet born. Unpublished paper, 1979.

Lunsford, A. Cognitive development and the basic writer. College English, 1977, 38, 38-46.

Lunsford, A. The content of basic writers' essays. Paper presented at the Annual Meeting of the Modern Language Association, San Francisco, 1979.

Maimon, E.P., and Nodine, B.F. Measuring syntactic growth: errors and expectations in sentence-combining practice with college freshmen. Research in the Teaching of English, 1978, 12, 233-244.

Marckworth, M.L., and Bell, L.M. Sentence length distribution in the corpus. In Kucera, H., and Francis, W.N. (Eds.), Computational analysis of present-day English. Providence, R.I.: Brown University Press, 1967.

Marzano, R.J. The sentence-combining myth. English Journal, 1976, 65, 57-59.

McCrimmon, J.M. Writing with a purpose. Boston: Houghton Mifflin, 1976.

Mellon, J. National assessment and the teaching of English. Urbana, IL: National Council of Teachers of English, 1975.

Mellon, J.C. Transformational sentence-combining: a method for enhancing the development of syntactic fluency in English composition. Research Report, No. 10. Urbana, Ill.: National Council of Teachers of English, 1969.

Meyer, B. The organization of prose and its effects on memory. Amsterdam: North Holland, 1975.

Meyer, B., and Freedle, R. Effects of discourse type on recall. Tempe: Arizona State University Department of Educational Psychology, Prose Learning Series, Report 6, 1979.

Moffett, J. Teaching the universe of discourse. Boston: Houghton Mifflin, 1968.

Morenberg, M., Daiber, D., and Kerek, A. Sentence combining at the college level: an experimental study. Research in the Teaching of English, 1978, 12, 245-256.

Mulder, J., Braun, C., and Holliday, W. Effects of sentence-combining practice on linguistic maturity levels of adult students. Adult Education, 1978, 28, 111-120.

Ney, J.W. The hazards of the course: sentence combining in freshman English. English Record, 1976, 27, 70-77.

Nold, E., and Freedman, S. An analysis of readers' responses to essays. Research in the Teaching of English, 1977, 11, 164-174.

O'Donnell, R.C. A critique of some indices of syntactic maturity. Research in the Teaching of English, 1976, 10, 31-38.

O'Donnell, R.C., Griffin, W., and Norris, R. *Syntax of kindergarten and elementary school children: a transformational analysis*. Research Report No. 8. Urbana, Ill.: National Council of Teachers of English, 1967.

O'Hare, F. *Sentence combining: improving student writing without formal grammar instruction*. Research Report No. 15. Urbana, Ill.: National Council of Teachers of English, 1973.

Perron, J.D. The impact of mode on written syntactic complexity: part I--third grade. *Studies in Language Education*, Report No. 24, Department of Language Education, University of Georgia, 1976a.

Perron, J.D. The impact of mode on written syntactic complexity: part II--fourth grade. *Studies in Language Education*, Report no. 25, Department of Language Education, University of Georgia, 1976b.

Perron, J.D. The impact of mode on written syntactic complexity: part III--fifth grade. Studies in Language Education, Report no. 26, Department of Language Education, University of Georgia, 1976c.

Perron, J.D. Written syntactic complexity and the modes of discourse. Paper presented during the Annual Meeting of the American Educational Research Association, New York, 1977. ERIC Document 139 009.

Pike, K.L. Beyond the sentence. *College Composition and Communication*, 1964, *15*, 129-135. b.

Pike, K.L. *Language in relation to a unified theory of the structure of human behavior*. The Hague: Mouton, 1967.

Pike, K.L. Language--where science and poetry meet. *College English*, 1965, *26*, 283-292.

Pike, K.L. A linguistic contribution to composition. *College Composition and Communication*, 1964, *15*, 82-88. a.

Pitkin, W. Discourse blocs. *College Composition and Communication*, 1969, *20*, 138-148.

Plutchik, R. Foundations of experimental research. New York: Harper & Row, 1974.

Raygor, A.L. McGraw-Hill Basic Skills System: Examiner's Manual. New York: McGraw-Hill 1970a.

Raygor, A.L. McGraw-Hill Basic Skills System Writing Test: Examiner's Manual. New York: McGraw-Hill, 1970b.

Richmond, V.P., and Dickson, F.C. Two studies of the validity of the writing apprehension test. Paper presented at the Annual Meeting of the Eastern Communication Association, 1980.

Rodgers, P. A discourse-centered rhetoric of the paragraph. College Composition and Communication, 1966, 17, 2-11.

Rodgers, P. The stadium of discourse. College Composition and Communication, 1967, 18, 178-185.

Rodi, S. Teaching syllabus for E306. Austin: University of Texas, 1978.

San Jose, C. Grammatical structures in four modes of writing at the fourth grade level. Unpublished doctoral dissertation, Syracuse University, 1972.

San Jose, C. Notes and comments. Research in the Teaching of English, 1978, 12, 91-92.

San Jose, C., and Remnick, B. The effects of stimuli on the grades and syntax of fifth graders writing in different modes. Unpublished paper, 1974.

Shaughnessy, M. Errors and expectations. New York: Oxford University Press, 1977.

Seegars, J.C. The form of discourse and sentence structure. Elementary English, 1933, 10, 51-54.

Sodowsky, R.E. and Witte, S.P. Positive measurement of the writing of college students. Paper presented during the Ve Congrès enternational de Linguistique Appliquée, 1978.

Stewart, M.F. Freshman sentence combining: a Canadian project. Research in the Teaching of English, 1978, 12, 257-268. a.

Stewart, M.F. Syntactic maturity from high school to university: a first look. *Research in the Teaching of English*, 1978, *12*, 37-46. b.

Strong, W. Sentence combining: a composing book. New York: Random House, 1973.

Thorndyke, P.C. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

Tibbetts, A.M. On the practical use of a grammatical system: a note on Christensen and Johnson. *College English*, 1970, *31*, 870-878.

Veal, L.R., and Tillman, M. Mode of discourse variation in the evaluation of children's writing. *Research in the Teaching of English*, 1971, *5*, 37-45.

Wesling, D., Conlisk, J., Evans, S., Hardison, W.G., Lovberg, R., Tolbert, E., and Watkins, J. *Evaluation of the four college writing programs at UC San Diego*. San Diego: University of California, San Diego, 1978.

Williams, J.M. Defining complexity, *College English*, 1979, *40*, 595-609.

Winterowd, W.R. *Contemporary rhetoric: a conceptual background with readings*. New York: Harcourt Brace Jovanovich, 1975.

Winterowd, W.R. The grammar of coherence. *College English*, 1970, *31*, 828-835.

Witte, S.P. Syntactic complexity and perceived writing quality: an issue in the teaching of freshman composition. Paper presented during the Annual Meeting of the Indiana Council of Teachers of English, 1978.

Witte, S.P. A neglected question in experimental and normative research in written composition: the stability of t-unit length. Paper presented during the Annual Meeting of the Conference on College Composition and Communication, 1979.

Witte, S.P., and Davis, A.S. The stability of T-unit length: a preliminary investigation. *Research in the Teaching of English*, 1980, *14*, 5-17.

Witte, S.P., and Davis, A.S. The stability of t-unit length in the written discourse of college freshmen: a second study. In Freedman, A., Pringle, I., and Yalden, J. (Eds.), Learning to Write. In press.

Witte, S.P., and Davis, W.J. Some tentative observations regarding the stability of t-unit length. Paper presented during the Annual Meeting of the Conference on College Composition and Communication, 1978.

Witte, S.P., and Sodowsky, R.E. A limited evaluation of OSU's experimental English composition program. Stillwater: Oklahoma State University, 1977.

Witte, S.P. and Sodowsky, R.E. Syntactic maturity in the writing of college freshmen. Paper presented during the Annual Meeting of the Conference on College Composition and Communication, 1978.

Wittig, S. STEPS to structure, Cambridge, Mass.: Winthrop 1975.

Young, R., and Becker, A. Toward a modern theory of rhetoric: a tagmemic contribution. Harvard Educational Review, 1965, 35, 450-468.